GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2003, 20:56:03 ; Search time 40 Seconds
(without alignments)
709.828 Million cell updates/sec

Title: US-09-677-752-2
Perfect score: 5086
Sequence: 1 MKKAFFFFLIGNSLSGLARE.....GFYSSSTFCNYLNGEIALRF 965

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4994.5	98.2	964	4	US-09-556-877-177	Sequence 177, App
2	4994.5	98.2	964	4	US-09-620-412C-177	Sequence 177, App
3	4919.5	96.7	977	4	US-09-556-877-191	Sequence 191, App
4	4919.5	96.7	977	4	US-09-620-412C-191	Sequence 191, App
5	647	12.7	982	4	US-09-556-877-176	Sequence 176, App
6	647	12.7	982	4	US-09-620-412C-176	Sequence 176, App
7	647	12.7	1006	4	US-09-556-877-190	Sequence 190, App
8	647	12.7	1006	4	US-09-620-412C-190	Sequence 190, App
9	605.5	11.9	715	4	US-09-620-412C-321	Sequence 321, App
10	471.5	9.3	880	4	US-09-556-877-175	Sequence 175, App
11	471.5	9.3	880	4	US-09-620-412C-175	Sequence 175, App
12	459.5	9.0	866	4	US-09-556-877-189	Sequence 189, App
13	459.5	9.0	866	4	US-09-620-412C-189	Sequence 189, App
14	398.5	7.8	1752	4	US-09-556-877-180	Sequence 180, App
15	398.5	7.8	1752	4	US-09-620-412C-180	Sequence 180, App
16	386.5	7.6	1530	4	US-09-556-877-178	Sequence 178, App
17	386.5	7.6	1530	4	US-09-620-412C-178	Sequence 178, App
18	371.5	7.3	715	4	US-09-620-412C-329	Sequence 329, App
19	365	7.2	848	4	US-09-556-877-192	Sequence 192, App
20	365	7.2	848	4	US-09-620-412C-192	Sequence 192, App
21	363	7.1	1776	4	US-09-556-877-179	Sequence 179, App
22	363	7.1	1776	4	US-09-620-412C-179	Sequence 179, App
23	345.5	6.8	948	4	US-09-556-877-194	Sequence 194, App
24	345.5	6.8	948	4	US-09-620-412C-194	Sequence 194, App
25	340	6.7	646	4	US-09-620-412C-317	Sequence 317, App
26	320.5	6.3	691	4	US-09-620-412C-313	Sequence 313, App
27	278.5	5.5	631	4	US-09-620-412C-325	Sequence 325, App
28	246.5	4.8	619	4	US-09-620-412C-309	Sequence 309, App
29	235	4.6	700	4	US-09-620-412C-345	Sequence 345, App
30	221.5	4.4	821	4	US-09-556-877-195	Sequence 195, App
31	221.5	4.4	821	4	US-09-620-412C-195	Sequence 195, App
32	220.5	4.3	654	4	US-09-620-412C-341	Sequence 341, App
33	215	4.2	778	4	US-09-556-877-193	Sequence 193, App
34	215	4.2	778	4	US-09-620-412C-193	Sequence 193, App
35	205	4.0	683	4	US-09-620-412C-357	Sequence 357, App
36	194	3.8	1612	1	US-08-169-927-2	Sequence 2, Appli
37	192.5	3.8	1095	4	US-09-206-942-45	Sequence 45, Appl
38	192.5	3.8	1101	4	US-09-206-942-43	Sequence 43, Appl
39	188.5	3.7	583	4	US-09-620-412C-353	Sequence 353, App
40	181.5	3.6	915	4	US-09-206-942-35	Sequence 35, Appl
41	181.5	3.6	1222	4	US-09-206-942-37	Sequence 37, Appl
42	181.5	3.6	1228	4	US-09-206-942-34	Sequence 34, Appl
43	177.5	3.5	1338	2	US-08-728-470-9	Sequence 9, Appli
44	177.5	3.5	1338	4	US-08-719-641-9	Sequence 9, Appli
45	177.5	3.5	1599	2	US-08-617-697-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-09-556-877-177
; Sequence 177, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 177
; LENGTH: 964
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-556-877-177

Query Match 98.2%; Score 4994.5; DB 4; Length 964;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 950; Conservative 7; Mismatches 7; Indels 1; Gaps 1;

```
Qy       1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 60

Qy      61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120

Qy     121 RDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVR 180
           |||||||||||||||  ||   || || :|||||||||||||||||||||||||||||||:
Db     121 RDTIGPVIFENNTCCRLFTWRNPYAA-DKIREGGAIHAQNLYINHNHDVVGFMKNFSYVQ 179

Qy     181 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 239

Qy     241 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     240 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 299

Qy     301 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFN 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     300 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFN 359
```

```
Qy    361 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 419

Qy    421 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 479

Qy    481 GVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNY 540
          |||||||||||||||| |:|||||||||||||||||||||||||||||||||||||||||||
Db    480 GVVSLGNGAVLSCYKNGTGDSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNY 539

Qy    541 TADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||:|:::||||||||
Db    540 TADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLQSENIDFSG 599

Qy    601 LNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRS 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 LNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRS 659

Qy    661 PLIANTLWGNMLLATESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSG 720
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    660 PLIANTLWGNMLLATESLKNSAELTPSGHPFWGITGGGLGMMVYQDPRENHPGFHMRSSG 719

Qy    721 YSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGL 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    720 YSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGL 779

Qy    781 YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYS 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    780 YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYS 839

Qy    841 SLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPG 900
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    840 SLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATHRPQAWTVELAYQPVLYRQEPG 899

Qy    901 IATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGE 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    900 IATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGE 959

Qy    961 IALRF 965
          |||||
Db    960 IALRF 964
```

RESULT 2
US-09-620-412C-177
; Sequence 177, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 177
; LENGTH: 964
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-620-412C-177

Query Match 98.2%; Score 4994.5; DB 4; Length 964;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 950; Conservative 7; Mismatches 7; Indels 1; Gaps 1;

```
Qy      1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 60
```

```
Qy     61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120

Qy    121 RDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVR 180
          |||||||||||||||| ||  || || :|||||||||||||||||||||||||||||||:
Db    121 RDTIGPVIFENNTCCRLFTWRNPYAA-DKIREGGAIHAQNLYINHNHDVVGFMKNFSYVQ 179

Qy    181 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 239

Qy    241 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 299

Qy    301 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFN 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFN 359

Qy    361 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 419

Qy    421 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 479

Qy    481 GVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNY 540
          |||||||||||||||| |:|||||||||||||||||||||||||||||||||||||||||||
Db    480 GVVSLGNGAVLSCYKNGTGDSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNY 539

Qy    541 TADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||:|:::||||||||
Db    540 TADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLQSENIDFSG 599

Qy    601 LNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRS 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 LNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRS 659

Qy    661 PLIANTLWGNMLLATESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSG 720
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    660 PLIANTLWGNMLLATESLKNSAELTPSGHPFWGITGGGLGMMVYQDPRENHPGFHMRSSG 719

Qy    721 YSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGL 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    720 YSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGL 779

Qy    781 YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYS 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    780 YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYS 839

Qy    841 SLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPG 900
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    840 SLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATHRPQAWTVELAYQPVLYRQEPG 899

Qy    901 IATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGE 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    900 IATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGE 959

Qy    961 IALRF 965
          |||||
Db    960 IALRF 964
```

RESULT 3
US-09-556-877-191
; Sequence 191, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir

```
;  APPLICANT:  Fling, Steve
;  APPLICANT:  Maisonneuve, Jeff
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE:  210121.469C5
;  CURRENT APPLICATION NUMBER: US/09/556,877
;  CURRENT FILING DATE:  2000-04-19
;  NUMBER OF SEQ ID NOS:  305
;  SOFTWARE:  FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 191
;   LENGTH: 977
;   TYPE: PRT
;   ORGANISM: Chlamydia
US-09-556-877-191

  Query Match             96.7%;  Score 4919.5;  DB 4;  Length 977;
  Best Local Similarity   98.4%;  Pred. No. 0;
  Matches  935;  Conservative    7;  Mismatches    7;  Indels    1;  Gaps    1;

Qy      16 GLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEG 75
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      29 GLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEG 88

Qy      76 AAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCC 135
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      89 AAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCC 148

Qy     136 RPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSEN 195
           | ||  || || :|||||||||||||||||||||||||||||:|||||||||||||||||
Db     149 RLFTWRNPYAA-DKIREGGAIHAQNLYINHNHDVVGFMKNFSYVQGGAISTANTFVVSEN 207

Qy     196 QSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNACCAGGAIFSPICSLTG 255
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     208 QSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNACCAGGAIFSPICSLTG 267

Qy     256 NRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYA 315
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     268 NRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYA 327

Qy     316 PVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTST 375
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     328 PVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTST 387

Qy     376 SANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVSFNKEADQTGSVVFSG 435
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     388 SANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVSFNKEADQTGSVVFSG 447

Qy     436 ATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYK 495
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     448 ATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYK 507

Qy     496 NGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKL 555
           || |:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     508 NGTGDSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKL 567

Qy     556 SLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSGLNVPHYGWQGLWTWG 615
           |||||||||||||||||||||||||||||||:|:::||||||||||||||||||||||||
Db     568 SLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLQSENIDFSGLNVPHYGWQGLWTWG 627

Qy     616 WAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLAT 675
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     628 WAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLAT 687

Qy     676 ESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFSL 735
           |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db     688 ESLKNSAELTPSGHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFSL 747

Qy     736 KFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQG 795
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     748 KFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQG 807

Qy     796 ENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSF 855
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     808 ENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSF 867

Qy     856 STKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGS 915
           |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db     868 STKTPLINVLVPIGVKGSFMNATHRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGS 927

Qy     916 GSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     928 GSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 977


RESULT 4
US-09-620-412C-191
; Sequence 191, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 191
;   LENGTH: 977
;   TYPE: PRT
;   ORGANISM: Chlamydia
US-09-620-412C-191

  Query Match             96.7%;  Score 4919.5;  DB 4;  Length 977;
  Best Local Similarity   98.4%;  Pred. No. 0;
  Matches  935;  Conservative    7;  Mismatches    7;  Indels    1;  Gaps    1;

Qy      16 GLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEG 75
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      29 GLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEG 88

Qy      76 AAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCC 135
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      89 AAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCC 148

Qy     136 RPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSEN 195
           | ||  || || :|||||||||||||||||||||||||||||:|||||||||||||||||
Db     149 RLFTWRNPYAA-DKIREGGAIHAQNLYINHNHDVVGFMKNFSYVQGGAISTANTFVVSEN 207

Qy     196 QSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNACCAGGAIFSPICSLTG 255
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     208 QSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNACCAGGAIFSPICSLTG 267

Qy     256 NRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYA 315
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     268 NRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYA 327

Qy     316 PVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTST 375
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     328 PVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTST 387

Qy     376 SANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVSFNKEADQTGSVVFSG 435
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     388 SANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVSFNKEADQTGSVVFSG 447

Qy     436 ATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYK 495
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     448 ATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYK 507

Qy     496 NGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKL 555
           || |:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     508 NGTGDSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKL 567

Qy     556 SLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSGLNVPHYGWQGLWTWG 615
           |||||||||||||||||||||||||||||||:|:::||||||||||||||||||||||||
Db     568 SLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLQSENIDFSGLNVPHYGWQGLWTWG 627
```

```
Qy    616 WAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLAT 675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    628 WAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLAT 687

Qy    676 ESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFSL 735
          ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    688 ESLKNSAELTPSGHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFSL 747

Qy    736 KFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQG 795
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    748 KFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQG 807

Qy    796 ENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSF 855
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    808 ENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSF 867

Qy    856 STKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGS 915
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    868 STKTPLINVLVPIGVKGSFMNATHRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGS 927

Qy    916 GSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    928 GSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 977
```

RESULT 5
US-09-556-877-176
; Sequence 176, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 176
; LENGTH: 982
; TYPE: PRT
; ORGANISM: Chlamydia
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(982)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-556-877-176

Query Match 12.7%; Score 647; DB 4; Length 982;
Best Local Similarity 25.4%; Pred. No. 2e-46;
Matches 257; Conservative 156; Mismatches 388; Indels 212; Gaps 43;

```
Qy     39 SLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAK 98
          :|    :: |  |:||    :|:|        |   |||::       ::  :|::
Db     54 NLLGSFTVLGRGHSLT---FENIR--------TSTNGAALS--------NSAADGLF--- 91

Qy     99 NLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNP-NAAV---------- 147
            | |    : ::: ||    :     |    |     | |:| | |  :
Db     92 --TIEGFKELSFSNCNSLLAVL-----PAATTNKGSQTPTTTSTPSNGTIYSKTDLLLLN 144

Qy    148 --------NKIR-EGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSC 198
                  | :  :|||| |::| :     :  | :| :   |||     :|    |::
Db    145 NEKFSFYSNLVSGDGGAIDAKSLTVQGISKLCVFQENTAQADGGACQVVTSFSAMANEAP 204

Qy    199 FLFMDNICIQTNTAG-KGGAIYA------GTSNSFE--------SNNCDLFFINNACCAG 243
            |: |:     || :|| | |      | |:|          | |  : |  |    |
Db    205 IAFVANV------AGVRGGGIAAVQDGQQGVSSSTSTEDPVVSFSRNTAVEFDGNVARVG 258

Qy    244 GAIFSPICSLTGNRGNIVFYNN--RCFKN----------------VETASSEASDGGAI 284
          | |:|        ||: | ||    | |                        |:   |||||
Db    259 GGIYS--------YGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQASNTSNNYGDGGAI 310

Qy    285 KVTTRLDVTGN---------RGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANN 335
                  |          | :||| |:   ||||||  ::: : ||  |: ||||:
Db    311 FCKNGAQAGSNNSGSVSFDGEGVVFFSSNVAAGKGGAIYAKKLSVANCGPVQFLRNIAND 370

Qy    336 KGGAIYIDGTSNSKISADRHAIIFNENI----VTNVTNANGTSTSANPPRRNAITVASSS 391
          |||||:  :    :|||   |||: |:       |  : || : |:      ||:: |
Db    371 -GGAIYLGESGELSLSADYGDIIFDGNLKRTAKENAADVNGVTVSS-----QAISMGSGG 424

Qy    392 GEILLGAGSSQNLIFYDPIEVSN-------AGVSVSFNKEADQTGSVVFSGATVNSADFH 444
              | :   ::| ||||::|       :   :  |     || :||:     |:  :
Db    425 KITTLRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIVFAN---GSSTLY 481

Qy    445 QRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNS--A 502
          |          :|:  | : : : |:|:|:|  :|||| : :  |: |            |
Db    482 QN---------VTIEQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTLDFVTPQPPQQPPA 532

Qy    503 SNASITLKHIGLNLSSILKSGAEIPLLWVEPTN--NSNNYTADTAATFSLSDVKLS---L 557
          :|  ||| :: |:|||:| : |    :   |||    ::: |   :| :   | :|
Db    533 ANQLITLSNLHLSLSSLLANNA----VTNPPTNPPAQDSHPAVIGST-TAGSVTISGPIF 587

Qy    558 IDDYGNSPYESTD----------LTHALSSQPMLSISEASDNQLRSDDMDFSGLNVPHYG 607
           :|  :: |:  |           |   | ::|    :  ||  |        |  :| ||
Db    588 FEDLDDTAYDRYDWLGSNQKINVLKLQLGTKP--PANAPSDLTL--------GNEMPKYG 637

Qy    608 WQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTL 667
          :|| |   |                ||  ||    ||  ||   || | |:  : |: |:|
Db    638 YQGSWKLAW--------------DPNTANNGPYTLKATWTKTGYNPGPERVASLVPNSL 682

Qy    668 WGNMLLATESLKNSAELTPSDHPFW-GITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMI 726
          ||: :|   |  :: : :    :  |:   |:    | |      |:   | ||| |
Db    683 WGS-ILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYHDRDALGQGYRYISGGYSLG-- 739

Qy    727 AGQTHTFSLKFSQTYTKLNERYAKNNV--SSKNYSCQGEMLFSLQEGFLLTKLVGLYSYG 784
             ::  |  |   :|::  | :|: |   | :::| | :  | |:      | | | :|
Db    740 -ANSYFGSSMFGLAFTEVFGR-SKDYVVCRSNHHACIGSVYLSTQQA-----LCGSYLFG 792

Qy    785 DHNCHHFYTQG-ENLTSQGTFRSQT--------MGGAVFFDLPMKPFGSTHILT--APFL 833
          |      |  | ::: :  ||  ::        : | :   ||:    |   |    ||:
Db    793 DAFIRASYGFGNQHMKTSYTFAEESDVRWDNNCLAGEIGAGLPIVITPSKLYLNELRPFV 852

Qy    834 GALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPV 893
          |   |:   ||| |   |:| :   |:|: ||:|||    ::|  |  ::   ||
Db    853 QAEFSYADHESFTEEGDQARAFKS-GHLLNLSVPVGVKFDRCSST-HPNKYSFMAAYICD 910

Qy    894 LYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHG 946
          ||   |  | ||: :  |       :|| :  : |        || : : :|
Db    911 AYRTISGTETTLLSHQETWTTDAFHLARHGVVVRGSMYAS----LTSNIEVYG 959
```

RESULT 6
US-09-620-412C-176
; Sequence 176, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 176
; LENGTH: 982
; TYPE: PRT
; ORGANISM: Chlamydia
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(982)

```
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-09-620-412C-176

  Query Match            12.7%;  Score 647;  DB 4;  Length 982;
  Best Local Similarity   25.4%;  Pred. No. 2e-46;
  Matches  257;  Conservative  156;  Mismatches  388;  Indels  212;  Gaps   43;

Qy      39 SLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAK 98
           :|    :: |  |:||    :|:|       |   |||::        ::  :|::
Db      54 NLLGSFTVLGRGHSLT---FENIR--------TSTNGAALS--------NSAADGLF--- 91

Qy      99 NLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNP-NAAV---------- 147
             | |    : ::: ||  :     |    |     | |:| | |  :
Db      92 --TIEGFKELSFSNCNSLLAVL-----PAATTNKGSQTPTTTSTPSNGTIYSKTDLLLLN 144

Qy     148 --------NKIR-EGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSC 198
                  | :  :|||| |::| :      :  | :| :   |||     :|    |::
Db     145 NEKFSFYSNLVSGDGGAIDAKSLTVQGISKLCVFQENTAQADGGACQVVTSFSAMANEAP 204

Qy     199 FLFMDNICIQTNTAG-KGGAIYA------GTSNSFE--------SNNCDLFFINNACCAG 243
             |: |:      || :|| | |      | |:|          | |  : |  |    |
Db     205 IAFVANV------AGVRGGGIAAVQDGQQGVSSSTSTEDPVVSFSRNTAVEFDGNVARVG 258

Qy     244 GAIFSPICSLTGNRGNIVFYNN--RCFKN-----------------VETASSEASDGGAI 284
           | |:|         ||: | ||    | |                      |:   |||||
Db     259 GGIYS--------YGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQASNTSNNYGDGGAI 310

Qy     285 KVTTRLDVTGN---------RGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANN 335
                  |           | :||| |:    ||||||  ::: : ||  |: ||||:
Db     311 FCKNGAQAGSNNSGSVSFDGEGVVFFSSNVAAGKGGAIYAKKLSVANCGPVQFLRNIAND 370

Qy     336 KGGAIYIDGTSNSKISADRHAIIFNENI----VTNVTNANGTSTSANPPRRNAITVASSS 391
           |||||:  :    :|||   |||: |:      |  : || : |:       ||:: |
Db     371 -GGAIYLGESGELSLSADYGDIIFDGNLKRTAKENAADVNGVTVSS-----QAISMGSGG 424

Qy     392 GEILLGAGSSQNLIFYDPIEVSN-------AGVSVSFNKEADQTGSVVFSGATVNSADFH 444
           | | :   ::| ||||::|       :   :  |     || :||:       |:  :
Db     425 KITTLRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIVFAN---GSSTLY 481

Qy     445 QRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNS--A 502
           |          :|:  | : : :  |:|:||  :|||| : :  |: |              |
Db     482 QN---------VTIEQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTLDFVTPQPPQQPPA 532

Qy     503 SNASITLKHIGLNLSSILKSGAEIPLLWVEPTN--NSNNYTADTAATFSLSDVKLS---L 557
           :|  ||| :: |:|||:| : |    :   |||    ::: |   :|  :   | :|
Db     533 ANQLITLSNLHLSLSSLLANNA----VTNPPTNPPAQDSHPAVIGST-TAGSVTISGPIF 587

Qy     558 IDDYGNSPYESTD----------LTHALSSQPMLSISEASDNQLRSDDMDFSGLNVPHYG 607
           :|  :: |:  |          |   | ::|    :  ||  |         |  :| ||
Db     588 FEDLDDTAYDRYDWLGSNQKINVLKLQLGTKP--PANAPSDLTL--------GNEMPKYG 637

Qy     608 WQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTL 667
           :|| |   |               || ||    ||  ||   || | |:  : |:|:|
Db     638 YQGSWKLAW--------------DPNTANNGPYTLKATWTKTGYNPGPERVASLVPNSL 682

Qy     668 WGNMLLATESLKNSAELTPSDHPFW-GITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMI 726
           ||: :|   |  :: : :     :  |:   |:    | |       |:   | |||  |
Db     683 WGS-ILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYHDRDALGQGYRYISGGYSLG-- 739

Qy     727 AGQTHTFSLKFSQTYTKLNERYAKNNV--SSKNYSCQGEMLFSLQEGFLLTKLVGLYSYG 784
             ::  |  |   :|::  | :|: |   | :::| | :  | |:      | | | :|
Db     740 -ANSYFGSSMFGLAFTEVFGR-SKDYVVCRSNHHACIGSVYLSTQQA-----LCGSYLFG 792

Qy     785 DHNCHHFYTQG-ENLTSQGTFRSQT--------MGGAVFFDLPMKPFGSTHILT--APFL 833
           |      |  | ::: :  ||  ::        : | :   ||:     |    |    ||:
Db     793 DAFIRASYGFGNQHMKTSYTFAEESDVRWDNNCLAGEIGAGLPIVITPSKLYLNELRPFV 852

Qy     834 GALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPV 893
           |   |:    ||| |   |:| :   |:|: ||:|||    ::|  |  ::   ||
Db     853 QAEFSYADHESFTEEGDQARAFKS-GHLLNLSVPVGVKFDRCSST-HPNKYSFMAAYICD 910

Qy     894 LYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHG 946
           ||   |  | ||: :  |       :|| :  : |        || : : :|
```

```
Db     911 AYRTISGTETTLLSHQETWTTDAFHLARHGVVVRGSMYAS----LTSNIEVYG 959


RESULT 7
US-09-556-877-190
; Sequence 190, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
;  APPLICANT:  Probst, Peter
;  APPLICANT:  Bhatia, Ajay
;  APPLICANT:  Skeiky, Yasir
;  APPLICANT:  Fling, Steve
;  APPLICANT:  Maisonneuve, Jeff
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE:  210121.469C5
;  CURRENT APPLICATION NUMBER: US/09/556,877
;  CURRENT FILING DATE:  2000-04-19
;  NUMBER OF SEQ ID NOS:  305
;  SOFTWARE:  FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 190
;   LENGTH: 1006
;   TYPE: PRT
;   ORGANISM: Chlamydia
US-09-556-877-190

  Query Match            12.7%;  Score 647;  DB 4;  Length 1006;
  Best Local Similarity   25.4%;  Pred. No. 2.1e-46;
  Matches  257;  Conservative  156;  Mismatches  388;  Indels  212;  Gaps   43;

Qy      39 SLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAK 98
           :|    :: |  |:||    :|:|       |   |||::        ::  :|::
Db      78 NLLGSFTVLGRGHSLT---FENIR--------TSTNGAALS--------NSAADGLF--- 115

Qy      99 NLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNP-NAAV---------- 147
             | |    : ::: ||    :     |    |     | |:| | |  :
Db     116 --TIEGFKELSFSNCNSLLAVL-----PAATTNKGSQTPTTTSTPSNGTIYSKTDLLLLN 168

Qy     148 --------NKIR-EGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSC 198
                  | :  :|||| |::| :      :  | :| :   |||     :|    |::
Db     169 NEKFSFYSNLVSGDGGAIDAKSLTVQGISKLCVFQENTAQADGGACQVVTSFSAMANEAP 228

Qy     199 FLFMDNICIQTNTAG-KGGAIYA------GTSNSFE--------SNNCDLFFINNACCAG 243
             |: |:      || :|| | |      | |:|          | |  : |  |    |
Db     229 IAFVANV------AGVRGGGIAAVQDGQQGVSSSTSTEDPVVSFSRNTAVEFDGNVARVG 282

Qy     244 GAIFSPICSLTGNRGNIVFYNN--RCFKN-----------------VETASSEASDGGAI 284
           | |:|         ||: | ||    | |                      |:   |||||
Db     283 GGIYS--------YGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQASNTSNNYGDGGAI 334

Qy     285 KVTTRLDVTGN---------RGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANN 335
                  |           | :||| |:    ||||||  ::: : ||  |: ||||:
Db     335 FCKNGAQAGSNNSGSVSFDGEGVVFFSSNVAAGKGGAIYAKKLSVANCGPVQFLRNIAND 394

Qy     336 KGGAIYIDGTSNSKISADRHAIIFNENI----VTNVTNANGTSTSANPPRRNAITVASSS 391
           |||||:  :    :|||   |||: |:      |  : || : |:       ||:: |
Db     395 -GGAIYLGESGELSLSADYGDIIFDGNLKRTAKENAADVNGVTVSS-----QAISMGSGG 448

Qy     392 GEILLGAGSSQNLIFYDPIEVSN-------AGVSVSFNKEADQTGSVVFSGATVNSADFH 444
           | | :   ::| ||||::|       :   :  |     || :||:       |:  :
Db     449 KITTLRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIVFAN---GSSTLY 505

Qy     445 QRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNS--A 502
           |          :|:  | : : :  |:|:||  :|||| : :  |: |              |
Db     506 QN---------VTIEQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTLDFVTPQPPQQPPA 556

Qy     503 SNASITLKHIGLNLSSILKSGAEIPLLWVEPTN--NSNNYTADTAATFSLSDVKLS---L 557
           :|  ||| :: |:|||:| : |    :   |||    ::: |   :|  :   | :|
Db     557 ANQLITLSNLHLSLSSLLANNA----VTNPPTNPPAQDSHPAVIGST-TAGSVTISGPIF 611

Qy     558 IDDYGNSPYESTD----------LTHALSSQPMLSISEASDNQLRSDDMDFSGLNVPHYG 607
           :|  :: |:  |          |   | ::|    :  ||  |         |  :| ||
Db     612 FEDLDDTAYDRYDWLGSNQKINVLKLQLGTKP--PANAPSDLTL--------GNEMPKYG 661
```

```
Qy    608 WQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTL 667
          :|| |   |                 ||  ||    ||  ||   || | |:  : |: |:|
Db    662 YQGSWKLAW--------------DPNTANNGPYTLKATWTKTGYNPGPERVASLVPNSL 706

Qy    668 WGNMLLATESLKNSAELTPSDHPFW-GITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMI 726
          ||: :|   |  :: : :     :  |:   |:    | |       |:   | ||| |
Db    707 WGS-ILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYHDRDALGQGYRYISGGYSLG-- 763

Qy    727 AGQTHTFSLKFSQTYTKLNERYAKNNV--SSKNYSCQGEMLFSLQEGFLLTKLVGLYSYG 784
            ::  |  |   :|::  | :|: |   | :::| | :  | |:      | | | :|
Db    764 -ANSYFGSSMFGLAFTEVFGR-SKDYVVCRSNHHACIGSVYLSTQQA-----LCGSYLFG 816

Qy    785 DHNCHHFYTQG-ENLTSQGTFRSQT--------MGGAVFFDLPMKPFGSTHILT--APFL 833
          |       |  | ::: :  ||  ::        : | :   ||:    |   |    ||:
Db    817 DAFIRASYGFGNQHMKTSYTFAEESDVRWDNNCLAGEIGAGLPIVITPSKLYLNELRPFV 876

Qy    834 GALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPV 893
            |   |:    ||| |   |:| :   |:|: ||:|||    ::|  |  ::   ||
Db    877 QAEFSYADHESFTEEGDQARAFKS-GHLLNLSVPVGVKFDRCSST-HPNKYSFMAAYICD 934

Qy    894 LYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHG 946
            ||   |  | ||: :  |       :|| :  : |        || : : :|
Db    935 AYRTISGTETTLLSHQETWTTDAFHLARHGVVVRGSMYAS----LTSNIEVYG 983
```

RESULT 8
US-09-620-412C-190
; Sequence 190, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 190
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-620-412C-190

Query Match 12.7%; Score 647; DB 4; Length 1006;
Best Local Similarity 25.4%; Pred. No. 2.1e-46;
Matches 257; Conservative 156; Mismatches 388; Indels 212; Gaps 43;

```
Qy     39 SLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAK 98
          :|    :: |  |:||    :|:|        |   |||::        ::  :|::
Db     78 NLLGSFTVLGRGHSLT---FENIR--------TSTNGAALS--------NSAADGLF--- 115

Qy     99 NLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNP-NAAV---------- 147
            | |    : ::: ||    :     |    |     | |:| | | :
Db    116 --TIEGFKELSFSNCNSLLAVL-----PAATTNKGSQTPTTTSTPSNGTIYSKTDLLLLN 168

Qy    148 --------NKIR-EGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSC 198
                 | :  :|||| |::| :     :  | :| :   |||     :|    |::
Db    169 NEKFSFYSNLVSGDGGAIDAKSLTVQGISKLCVFQENTAQADGGACQVVTSFSAMANEAP 228

Qy    199 FLFMDNICIQTNTAG-KGGAIYA------GTSNSFE--------SNNCDLFFINNACCAG 243
            |: |:      || :|| | |      | |:|          | |  : |  |    |
Db    229 IAFVANV------AGVRGGGIAAVQDGQQGVSSSTSTEDPVVSFSRNTAVEFDGNVARVG 282

Qy    244 GAIFSPICSLTGNRGNIVFYNN--RCFKN----------------VETASSEASDGGAI 284
          | |:|         ||: | ||    | |                  |:   |||||
Db    283 GGIYS--------YGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQASNTSNNYGDGGAI 334

Qy    285 KVTTRLDVTGN---------RGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANN 335
                  |           | :||| |:    ||||||  ::: : ||  |: ||||:
Db    335 FCKNGAQAGSNNSGSVSFDGEGVVFFSSNVAAGKGGAIYAKKLSVANCGPVQFLRNIAND 394

Qy    336 KGGAIYIDGTSNSKISADRHAIIFNENI----VTNVTNANGTSTSANPPRRNAITVASSS 391
          |||||:  :    :|||   |||: |: .     |  : || : |:      ||:: |
Db    395 -GGAIYLGESGELSLSADYGDIIFDGNLKRTAKENAADVNGVTVSS-----QAISMGSGG 448

Qy    392 GEILLGAGSSQNLIFYDPIEVSN-------AGVSVSFNKEADQTGSVVFSGATVNSADFH 444
            | | :   ::| ||||::|        :   :  |      || :||:      |:  :
Db    449 KITTLRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIVFAN---GSSTLY 505

Qy    445 QRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNS--A 502
          |          :|:  | : : :  |:|:||  :|||| : :  |: |             |
Db    506 QN---------VTIEQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTLDFVTPQPPQQPPA 556

Qy    503 SNASITLKHIGLNLSSILKSGAEIPLLWVEPTN--NSNNYTADTAATFSLSDVKLS---L 557
          :|  ||| :: |:|||:| : |    :   |||    ::: |   :| :   | :|
Db    557 ANQLITLSNLHLSLSSLLANNA----VTNPPTNPPAQDSHPAVIGST-TAGSVTISGPIF 611

Qy    558 IDDYGNSPYESTD----------LTHALSSQPMLSISEASDNQLRSDDMDFSGLNVPHYG 607
          :|  :: |:  |          |  | ::|    :  ||  |         |  :| ||
Db    612 FEDLDDTAYDRYDWLGSNQKINVLKLQLGTKP--PANAPSDLTL--------GNEMPKYG 661

Qy    608 WQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTL 667
          :|| |   |                 ||  ||    ||  ||   || | |:  : |: |:|
Db    662 YQGSWKLAW--------------DPNTANNGPYTLKATWTKTGYNPGPERVASLVPNSL 706

Qy    668 WGNMLLATESLKNSAELTPSDHPFW-GITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMI 726
          ||: :|   |  :: : :     :  |:   |:    | |       |:   | ||| |
Db    707 WGS-ILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYHDRDALGQGYRYISGGYSLG-- 763

Qy    727 AGQTHTFSLKFSQTYTKLNERYAKNNV--SSKNYSCQGEMLFSLQEGFLLTKLVGLYSYG 784
            ::  |  |   :|::  | :|: |   | :::| | :  | |:      | | | :|
Db    764 -ANSYFGSSMFGLAFTEVFGR-SKDYVVCRSNHHACIGSVYLSTQQA-----LCGSYLFG 816

Qy    785 DHNCHHFYTQG-ENLTSQGTFRSQT--------MGGAVFFDLPMKPFGSTHILT--APFL 833
          |       |  | ::: :  ||  ::        : | :   ||:    |   |    ||:
Db    817 DAFIRASYGFGNQHMKTSYTFAEESDVRWDNNCLAGEIGAGLPIVITPSKLYLNELRPFV 876

Qy    834 GALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPV 893
            |   |:    ||| |   |:| :   |:|: ||:|||    ::|  |  ::   ||
Db    877 QAEFSYADHESFTEEGDQARAFKS-GHLLNLSVPVGVKFDRCSST-HPNKYSFMAAYICD 934

Qy    894 LYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHG 946
            ||   |  | ||: :  |       :|| :  : |        || : : :|
Db    935 AYRTISGTETTLLSHQETWTTDAFHLARHGVVVRGSMYAS----LTSNIEVYG 983
```

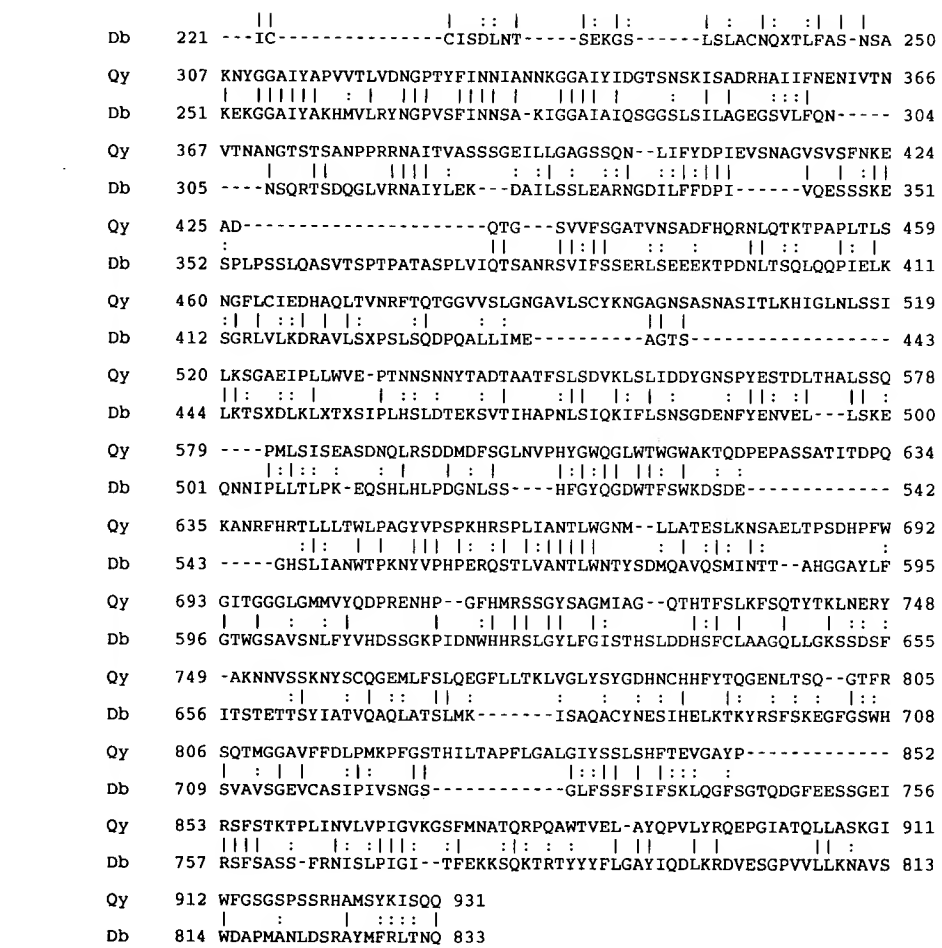
RESULT 9
US-09-620-412C-321
; Sequence 321, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 321
; LENGTH: 715
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-09-620-412C-321

Query Match 11.9%; Score 605.5; DB 4; Length 715;
Best Local Similarity 27.2%; Pred. No. 4.4e-43;
Matches 208; Conservative 120; Mismatches 277; Indels 159; Gaps 30;

```
Qy    302 SDNITKNYGGAIYA-PV-----------VTLVDNGPTYFIN-NIANNKGGAIYIDGTSNS 348
          |||   : ||  :| |:           :  |  ||| |:   : :| |    :      |
Db     11 SDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRVVGS 70

Qy    349 KISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYD 408
```

```
Db     71 APAASLG--ISTGDVITAVDGA-------------PINSATAMADALNG---------HH 106

Qy    409 PIEVSNAGVSVSFNKEA--DQTGSVVFSGATVNSADF-----HQRNLQTKT--------- 452
Db    107 PGDV----ISVTWQTKSGGTRTGNVTL--AEGPPAEFCRYPSHWRPLDIRTLMGKEHNYI 160

Qy    453 -PAPLTLSNGFLCIEDHAQLTVNR--FTQT-GGVVSLGNGAVLSCYKNGAGNSASNASIT 508
Db    161 KEAPTTLKFGTLAIEDDAELEIFNIPFTQNPTSLLALGSGATLTVGKHG--------KLN 212

Qy    509 LKHIGLNLSSILKSGAEIPLLWVEPTN-NSNNYTADT-AATFSLS------DVKLSLIDD 560
Db    213 ITNLGVILPIILKEGKSPPCIRVNPQDMTQNTGTGQTPSSTSSISTPMIIFNGRLSIVDE 272

Qy    561 YGNSPYESTDLTHALSSQPMLSISEASDNQLRSD---DMDFSGLNVPHYGWQGLWTWGWA 617
Db    273 NYESVYDSMDLSRGKAEQLILSIETTNDGQLDSNWQSSLNTSLLSPPHYGYQGLWTPNWI 332

Qy    618 KT--------QDPEPASSATITDPQKANR------------------------------ 638
Db    333 TTTYTITLNNNSSAPTSATSIAEQKKTSETFTPSNTTTASIPNIKASAGSGSGSASNSGE 392

Qy    639 ---FHRTLLLTWLPAGYVPSPKHRSPLIANTL---WGNMLLATESLKNSAELTPSDHPFW 692
Db    393 VTITKHTLVVNWAPVGYIVDPIRRGDLIANSLVHSGRNMTMGLRSLL-------PDNSWF 445

Qy    693 GITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIA--GQTHTFSLKFSQTYTKLNERYAK 750
Db    446 ALQGAATTLFTKQQKRLSYHGYSSASKGYTVSSQASGAHGHKFLLSFSQSSDKMKEKETN 505

Qy    751 NNVSSKNY---SCQGEMLFSLQEGFLLTKLVG--LYSYGDHNCHHFYTQGENLTSQGTFR 805
Db    506 NRLSSRYYLSALCFEHPMFD------RIALIGAAACNYGTHNMRSFY--GTKKSSKGKFH 557

Qy    806 SQTMGGAVFFDL----PMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPL 861
Db    558 STTLGASLRCELRDSMPLR-----SIMLTPFAQALFSRTEPASIRESGDLARLFTLEQAH 612

Qy    862 INVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSSR 921
Db    613 TAVVSPIGIKGAYSSDTWPTLSWEMELAYQPTLYWKRPLLNTLLIQNNGSWVTTNTPLAK 672

Qy    922 HAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
Db    673 HSF-YGRGSHSLKFSHLKLFANYQAEVATSTVSHYINAGGALVF 715


RESULT 10
US-09-556-877-175
; Sequence 175, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
;  APPLICANT:  Probst, Peter
;  APPLICANT:  Bhatia, Ajay
;  APPLICANT:  Skeiky, Yasir
;  APPLICANT:  Fling, Steve
;  APPLICANT:  Maisonneuve, Jeff
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE:  210121.469C5
;  CURRENT APPLICATION NUMBER: US/09/556,877
;  CURRENT FILING DATE:  2000-04-19
;  NUMBER OF SEQ ID NOS:  305
;  SOFTWARE:  FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 175
;   LENGTH: 880
;   TYPE: PRT
;   ORGANISM: Chlamydia
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (1)...(880)
;   OTHER INFORMATION: Xaa = Any Amino Acid

US-09-556-877-175

  Query Match             9.3%;  Score 471.5;  DB 4;  Length 880;
  Best Local Similarity   23.5%;  Pred. No. 1.9e-31;
  Matches  231;  Conservative  147;  Mismatches  349;  Indels  257;  Gaps   46;

Qy     29 PNSV-----PDPTKESL----------SNKISLTGDTHNLTNCYLDNLRYILAILQKTPN 73
Db     40 PNHVVCTFFEDCTMESLFPALCAHASQDDPLYVLGNSY----CWFVSKLHITDPKEALFK 95

Qy     74 EGAAVTITD--YLSFFD-TQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFE 130
Db     96 EKGDLSIQNFRFLSFTDCSSKE--------------------SSPSI-IHQKNGQLSLR 133

Qy    131 NN---TCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTA 187
Db    134 NNGSMSFCRNHAEGS----------GGAISADAFSLQHNYLFTAFEENSSKGNGGAIQ-A 182

Qy    188 NTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCD-LFFINNACCAGGAI 246
Db    183 QTFSLSRNVSPISFARN-----RADLNGGAICC--SNLICSGNVNPLFFTGNSATNGGA- 234

Qy    247 FSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNIT 306
Db    235 ---IC---------------CISDLNT-----SEKGSLSLACNQET-------LFASNSA 264

Qy    307 KNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFNENIVTN 366
Db    265 KEKGGAIYAKHMVLRYNGPVSFINNSA-KIGGAIAIQSGGSLSILAGEGSVLFQN----- 318

Qy    367 VTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQN--LIFYDPIEVSNAGVSVSFNKE 424
Db    319 ----NSQRTSDQGLVRNAIYLXK---DAILSSLEARNGDILFFDPI------VQESSSKE 365

Qy    425 AD----------------------QTG---SVVFSGATVNSADFHQRNLQTKTPAPLTLS 459
Db    366 SPLPSSLQASVTSPTPATASPLVIQTSANRSVIFSSERLSEEEKTPDNLTSQLQQPIELK 425

Qy    460 NGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSI 519
Db    426 SGRLVLKDRAVLSAPSLSQDPQALLIME----------AGTSLKTSS------DLKLATL 469

Qy    520 LKSGAEIPLLWVEPTNNSNNYTADTAATF-----SLSDVKLSLIDDYGNSPYESTDLTHA 574
Db    470 -----SIPL-----------HSLDTEKSVTIHAPNLSIQKIFLSNSGDENFYENVEL--- 510

Qy    575 LSSQ----PMLSISEASDNQLRSDDMDFSGLNVPHYGWQGLWTWGWAKTQDPEPASSATI 630
Db    511 LSKEQNNIPLLTLPK-EQSHLHLPDGNLSS----HFGYQGDWTFSWKDSDE--------- 556

Qy    631 TDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNM--LLATESLKNSAELTPSD 688
Db    557 ---------GHSLIANWTPKNYVPHPERQSTLVANTLWNTYSDMQAVQSMINTT--AHGG 605

Qy    689 HPFWGITGGGLGMMVYQDPRENHP--GFHMRSSGYSAGMIAG--QTHTFSLKFSQTYTKL 744
Db    606 AYLFGTWGSAVSNLFYVHDSSGKPIDNWHHRSLGYLFGISTHSLDDHSFCLAAGQLLGKS 665

Qy    745 NERY-AKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQGENLTSQ-- 801
Db    666 SDSFITSTETTSYIATVQAQLATSLMK-------ISAQACYNESIHELKTKYRSFSKEGF 718

Qy    802 GTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYP--------- 852
Db    719 GSWHSVAVSGEVCASIPIVSNGS------------GLFSSFSIFSKLQGFSGTQDGFEES 766

Qy    853 ----RSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVEL-AYQPVLYRQEPGIATQLLA 907
Db    767 SGEIRSFSASS-FRNISLPIGI--TFEKKSQKTRTYYYFLGAYIQDLKRDVESGPVVLLK 823

Qy    908 SKGIWFGSGSPSSRHAMSYKISQQ 931
Db    824 NAVSWDAPMANLDSRAYMFRLTNQ 847
```

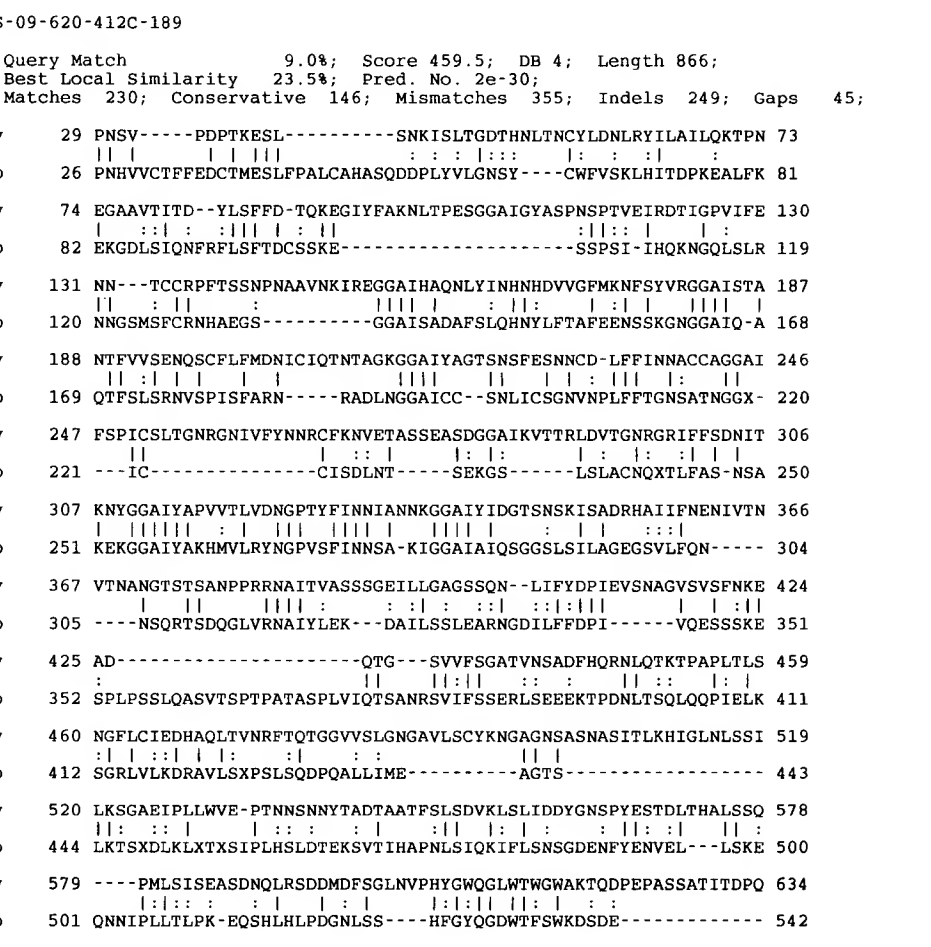
```
RESULT 11
US-09-620-412C-175
; Sequence 175, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 175
;   LENGTH: 880
;   TYPE: PRT
;   ORGANISM: Chlamydia
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (1)...(880)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-09-620-412C-175

  Query Match             9.3%;  Score 471.5;  DB 4;  Length 880;
  Best Local Similarity   23.5%;  Pred. No. 1.9e-31;
  Matches  231;  Conservative  147;  Mismatches  349;  Indels  257;  Gaps   46;

Qy      29 PNSV-----PDPTKESL----------SNKISLTGDTHNLTNCYLDNLRYILAILQKTPN 73
           || |       | | |||           : : : |:::    |:  :  :|    :
Db      40 PNHVVCTFFEDCTMESLFPALCAHASQDDPLYVLGNSY----CWFVSKLHITDPKEALFK 95

Qy      74 EGAAVTITD--YLSFFD-TQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFE 130
           |   ::| :  :||| | : ||                     :||:: |    | :
Db      96 EKGDLSIQNFRFLSFTDCSSKE--------------------SSPSI-IHQKNGQLSLR 133

Qy     131 NN---TCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTA 187
           ||   : ||     :          |||| |    : ||:    | :| |   |||| |
Db     134 NNGSMSFCRNHAEGS----------GGAISADAFSLQHNYLFTAFEENSSKGNGGAIQ-A 182

Qy     188 NTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCD-LFFINNACCAGGAI 246
           || :| | |   |  |          ||||    ||   | | : |||  |:   |||
Db     183 QTFSLSRNVSPISFARN-----RADLNGGAICC--SNLICSGNVNPLFFTGNSATNGGA- 234

Qy     247 FSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNIT 306
              ||               |  :: |     |: |:: :    :         |: |
Db     235 ---IC--------------CISDLNT-----SEKGSLSLACNQET-------LFASNSA 264

Qy     307 KNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFNENIVTN 366
           |  ||||||  : |  |||  |||| |   |||| |    :  | |   :::|
Db     265 KEKGGAIYAKHMVLRYNGPVSFINNSA-KIGGAIAIQSGGSLSILAGEGSVLFQN----- 318

Qy     367 VTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQN--LIFYDPIEVSNAGVSVSFNKE 424
           |   ||     |||| :     : :| :  ::|  ::|:|||      |  | :||
Db     319 ----NSQRTSDQGLVRNAIYLXK---DAILSSLEARNGDILFFDPI------VQESSSKE 365

Qy     425 AD---------------------QTG---SVVFSGATVNSADFHQRNLQTKTPAPLTLS 459
           :                      ||    ||:||   ::  :    || ::   |: |
Db     366 SPLPSSLQASVTSPTPATASPLVIQTSANRSVIFSSERLSEEEKTPDNLTSQLQQPIELK 425

Qy     460 NGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSI 519
           :| | ::| | |:    :|    : :             || |   :|       | |:::
Db     426 SGRLVLKDRAVLSAPSLSQDPQALLIME----------AGTSLKTSS------DLKLATL 469

Qy     520 LKSGAEIPLLWVEPTNNSNNYTADTAATF-----SLSDVKLSLIDDYGNSPYESTDLTHA 574
                |||           :: ||  :      :||  |: | :    : ||: :|
Db     470 -----SIPL-----------HSLDTEKSVTIHAPNLSIQKIFLSNSGDENFYENVEL--- 510

Qy     575 LSSQ----PMLSISEASDNQLRSDDMDFSGLNVPHYGWQGLWTWGWAKTQDPEPASSATI 630
           || :    |:|:: :   : |   | : |     |:|:|| ||: |  : :
Db     511 LSKEQNNIPLLTLPK-EQSHLHLPDGNLSS----HFGYQGDWTFSWKDSDE--------- 556

Qy     631 TDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNM--LLATESLKNSAELTPSD 688
                     :|:  | |  ||| |: :| |:|||||     : | :|: |:
Db     557 ---------GHSLIANWTPKNYVPHPERQSTLVANTLWNTYSDMQAVQSMINTT--AHGG 605

Qy     689 HPFWGITGGGLGMMVYQDPRENHP--GFHMRSSGYSAGMIAG--QTHTFSLKFSQTYTKL 744
           :|  |  :  : |       |   :| || ||  |:       |:| |   |   |
Db     606 AYLFGTWGSAVSNLFYVHDSSGKPIDNWHHRSLGYLFGISTHSLDDHSFCLAAGQLLGKS 665

Qy     745 NERY-AKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQGENLTSQ-- 801
           :: :      :|   : | ::  || :       :   :  : : |   |:  : : :
Db     666 SDSFITSTETTSYIATVQAQLATSLMK-------ISAQACYNESIHELKTKYRSFSKEGF 718

Qy     802 GTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYP--------- 852
           |:: |  : | |   :|:   ||            |::|| | |:::  :
Db     719 GSWHSVAVSGEVCASIPIVSNGS-----------GLFSSFSIFSKLQGFSGTQDGFEES 766

Qy     853 ----RSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVEL-AYQPVLYRQEPGIATQLLA 907
                |||| :   |: :|||:  :|   :|: : :   | ||   | |       ||
Db     767 SGEIRSFSASS-FRNISLPIGI--TFEKKSQKTRTYYYFLGAYIQDLKRDVESGPVVLLK 823

Qy     908 SKGIWFGSGSPSSRHAMSYKISQQ 931
           :   |    :     |  :::: |
Db     824 NAVSWDAPMANLDSRAYMFRLTNQ 847


RESULT 12
US-09-556-877-189
; Sequence 189, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
;  APPLICANT:  Probst, Peter
;  APPLICANT:  Bhatia, Ajay
;  APPLICANT:  Skeiky, Yasir
;  APPLICANT:  Fling, Steve
;  APPLICANT:  Maisonneuve, Jeff
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE:  210121.469C5
;  CURRENT APPLICATION NUMBER: US/09/556,877
;  CURRENT FILING DATE:  2000-04-19
;  NUMBER OF SEQ ID NOS:  305
;  SOFTWARE:  FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 189
;   LENGTH: 866
;   TYPE: PRT
;   ORGANISM: Chlamydia
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (1)...(866)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-09-556-877-189

  Query Match             9.0%;  Score 459.5;  DB 4;  Length 866;
  Best Local Similarity   23.5%;  Pred. No. 2e-30;
  Matches  230;  Conservative  146;  Mismatches  355;  Indels  249;  Gaps   45;

Qy      29 PNSV-----PDPTKESL----------SNKISLTGDTHNLTNCYLDNLRYILAILQKTPN 73
           || |       | | |||           : : : |:::    |:  :  :|    :
Db      26 PNHVVCTFFEDCTMESLFPALCAHASQDDPLYVLGNSY----CWFVSKLHITDPKEALFK 81

Qy      74 EGAAVTITD--YLSFFD-TQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFE 130
           |   ::| :  :||| | : ||                     :||:: |    | :
Db      82 EKGDLSIQNFRFLSFTDCSSKE--------------------SSPSI-IHQKNGQLSLR 119

Qy     131 NN---TCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTA 187
           ||   : ||     :          |||| |    : ||:    | :| |   |||| |
Db     120 NNGSMSFCRNHAEGS----------GGAISADAFSLQHNYLFTAFEENSSKGNGGAIQ-A 168

Qy     188 NTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCD-LFFINNACCAGGAI 246
           || :| | |   |  |          ||||    ||   | | : |||  |:   ||
Db     169 QTFSLSRNVSPISFARN-----RADLNGGAICC--SNLICSGNVNPLFFTGNSATNGGX- 220

Qy     247 FSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNIT 306
```



RESULT 13
US-09-620-412C-189
; Sequence 189, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 189
; LENGTH: 866
; TYPE: PRT
; ORGANISM: Chlamydia
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(866)
; OTHER INFORMATION: Xaa = Any Amino Acid

US-09-620-412C-189

Query Match 9.0%; Score 459.5; DB 4; Length 866;
Best Local Similarity 23.5%; Pred. No. 2e-30;
Matches 230; Conservative 146; Mismatches 355; Indels 249; Gaps 45;



```
RESULT 14
US-09-556-877-180
; Sequence 180, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
;  APPLICANT:  Probst, Peter
;  APPLICANT:  Bhatia, Ajay
;  APPLICANT:  Skeiky, Yasir
;  APPLICANT:  Fling, Steve
;  APPLICANT:  Maisonneuve, Jeff
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE:  210121.469C5
;  CURRENT APPLICATION NUMBER: US/09/556,877
;  CURRENT FILING DATE:  2000-04-19
;  NUMBER OF SEQ ID NOS:  305
;  SOFTWARE:  FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 180
;   LENGTH: 1752
;   TYPE: PRT
;   ORGANISM: Chlamydia
US-09-556-877-180

  Query Match             7.8%;  Score 398.5;  DB 4;  Length 1752;
  Best Local Similarity   23.8%;  Pred. No. 1.1e-24;
  Matches  271;  Conservative  135;  Mismatches  379;  Indels  353;  Gaps   60;

Qy      29 PNSVPDPTKESLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEGAAVT--------- 79
           | | |  |  :||  | |          : :|  : |  | :|      |
Db     653 PKSAPVST--ALSTPSSST----------VSSLTLLAASSQASPATSNKETQDPNADTDL 700

Qy      80 ITDYLSFFDTQKEGIYFAKNLTPESGGAIGYA------------SPNSPTVEIRDTIGP 126
           : ||:   ||       :|| | : || | ||             | || |      ||
Db     701 LIDYV--VDTT-----ISKN-TAKKGGGI-YAKKAKMSRIDQLNISENSAT-----EIG- 745

Qy     127 VIFENNTCCRPFTSSNPNAAV------NKI-REGGAIHAQNLYINHNHDVVGFMKN---- 175
                 ||:   |   :| |      | : :||| :||: : |::      |  |
Db     746 ----GGICCK--ESLELDALVSLSVTENLVGKEGGGLHAKTVNISNLKSGFSFSNNKANS 799

Qy     176 ---------------------------------FSYVRGGAI--------STANTFVVS 193
                                             :| | ||||          : |   |
Db     800 SSTGVATTASAPAAAAASLQAAAAAAPSSPATPTYSGVVGGAIYGEKVTFSQCSGTCQFS 859

Qy     194 ENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFI------------NNACC 241
           ||:    :||   |::   :||||| || |  |::    :|
Db     860 GNQA----IDNNPSQSSLNVQGGAIYAKTSLSIGSSDAGTSYIFSGNSVSTGKSQTTGQI 915

Qy     242 AGGAIFSPICSLTGNRGNIVFYNNRC-FKNVETASSEASDGGAIKVT-------TRLDVT 293
           |||||:||  :|        | ||       :|:| : | | :|| |        | : ::
Db     916 AGGAIYSPTVTL---NCPATFSNNTASIATPKTSSEDGSSGNSIKDTIGGAIAGTAITLS 972

Qy     294 G---------------------------------------NRGRIFFSDNITKNYGGAI 313
           |                                       |   | |   |  |||
Db     973 GVSRFSGNTADLGAAIGTLANANTPSATSGSQNSITEKITLENGSFIFERN-QANKRGAI 1031

Qy     314 YAPVVTLVDNGPTYFINNIANNKGGAIYI--DGTSNSKISADRHAIIFNENIVTNVTNAN 371
           |:| |::  |  | |  | : : | |||   | |  |  |     ::|  | ||  | |:
Db    1032 YSPSVSIKGNNIT-FNQNTSTHDGSAIYFTKDATIESLGS-----VLFTGNNVT-ATQAS 1084

Qy     372 GTSTSANPPRRN------------------AITVASSSGEILLGAGSSQN--------- 403
             ::  |    |                   :|: :||| |     | ||
Db    1085 SATSGQNTNTANYGAAIFGDPGTTQSSQTDAILTLLASSGNITFSNNSLQNNQGDTPASK 1144

Qy     404 -------------------LIFYDPIEVS--------NAGVSVSFNKEADQ---TGSVVF 433
                             : |:| :  |         |   ::  ||| :    ||::||
Db    1145 FCSIAGYVKLSLQAAKGKTISFFDCVHTSTKKTGSTQNVYETLDINKEENSNPYTGTIVF 1204

Qy     434 SGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQ-TGGVVSLGNGAVLS 492
                 |:: |:   ::  |    | || | :::  :| |  | |  |  : :  ||||
Db    1205 ------SSELHEN--KSYIPQNAILHNGTLVLKEKTELHVVSFEQKEGSKLIMEPGAVLS 1256

Qy     493 CYKNGAGNSASNASITLKHIGLNLSSI-LKSGAEI---PLLWVEPTNNSNN--------- 539
                   : :| :: :  : ::|||:      ||   | | :  | :| :
Db    1257 ------NQNIANGALAINGLTIDLSSMGTPQAGEIFSPPELRIVATTSSASGGSGVSSSI 1310

Qy     540 ----------YTADTAATFSLSDVK------LSLIDDYGN---SPYESTDLTHALSSQP 579
                      :   | | ::|: |      |:||| ||   :|   :||   |   |
Db    1311 PTNPKRISAAVPSGSAATTPTMSENKVFLTGDLTLIDPNGNFYQNPMLGSDLDVPLIKLP 1370

Qy     580 MLSISEASDNQLRSDDMDFSGLNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRF 639
               :  || |:   |:  ||   |  |: | ||                  ::||
Db    1371 ----TNTSDVQVY--DLTLSGDLFPQKGYMGTWTLD--------------SNPQTGKLQ 1409

Qy     640 HRTLLLTWLPAGYVPSPKHRSPLIANTLWG---NMLLATESL----KNSAELTP-SDHPF 691
           |    |:    |:|   |      ||:: |   :|::  : |     |:|     : : |
Db    1410 ARWTFDTYRRWVYIPRDNH---FYANSILGSQNSMIVVKQGLINNMLNNARFDDIAYNNF 1466

Qy     692 WGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFSL--KFSQTYTKLNERYA 749
           |     |:|  : |         |   | | |  : |      | |   ||:   |
Db    1467 W---VSGVGTFLAQQGTPLSEEFSYYSRGTSVAIDAKPRQDFILGAAFSKIVGKTKAIKK 1523

Qy     750 KNNVSSK--NYSCQ-----GEMLFSL---QEGFLLTKLV-GLYSYGDHNCHHFYTQGENL 798
            :|   |   || |      |: |: |   | |: |  |: |: ||| |  |   |   ::
Db    1524 MHNYFHKGSEYSYQASVYGGKFLYFLLNKQHGWALPFLIQGVVSYG-HIKHDTTTLYPSI 1582

Qy     799 TSQGTFRSQTMGG----AVFFDLPMKPFGSTHILTAPFLGALGIYSSL--SHFTEVGAYP 852
              :    : :|      :  ||       |:  :|    | |  |||:    |||:   |
Db    1583 HERNKGDWEDLGWLADLRISMDLKEPSKDSSKRITV--YGELE-YSSIRQKQFTEIDYDP 1639

Qy     853 RSFSTKTPLINVLVPIG--VKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLAS 908
           | |       |: :|:|  |:|: ||         : ||| | :||  |    ::|:|
Db    1640 RHFD-DCAYRNLSLPVGCAVEGAIMNCNILMYN-KLALAYMPSIYRNNPVCKYRVLSS 1695


RESULT 15
US-09-620-412C-180
; Sequence 180, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 180
;   LENGTH: 1752
;   TYPE: PRT
;   ORGANISM: Chlamydia
US-09-620-412C-180

  Query Match             7.8%;  Score 398.5;  DB 4;  Length 1752;
  Best Local Similarity   23.8%;  Pred. No. 1.1e-24;
  Matches  271;  Conservative  135;  Mismatches  379;  Indels  353;  Gaps   60;

Qy      29 PNSVPDPTKESLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEGAAVT--------- 79
           | | |  |  :||  | |          : :|  : |  | :|      |
Db     653 PKSAPVST--ALSTPSSST----------VSSLTLLAASSQASPATSNKETQDPNADTDL 700

Qy      80 ITDYLSFFDTQKEGIYFAKNLTPESGGAIGYA------------SPNSPTVEIRDTIGP 126
           : ||:   ||       :|| | : || | ||             | || |      ||
Db     701 LIDYV--VDTT-----ISKN-TAKKGGGI-YAKKAKMSRIDQLNISENSAT-----EIG- 745

Qy     127 VIFENNTCCRPFTSSNPNAAV------NKI-REGGAIHAQNLYINHNHDVVGFMKN---- 175
                 ||:   |   :| |      | : :||| :||: : |::      |  |
Db     746 ----GGICCK--ESLELDALVSLSVTENLVGKEGGGLHAKTVNISNLKSGFSFSNNKANS 799

Qy     176 ---------------------------------FSYVRGGAI--------STANTFVVS 193
                                             :| | ||||          : |   |
Db     800 SSTGVATTASAPAAAAASLQAAAAAAPSSPATPTYSGVVGGAIYGEKVTFSQCSGTCQFS 859

Qy     194 ENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFI------------NNACC 241
```

```
Db    860 GNQA----IDNNPSQSSLNVQGGAIYAKTSLSIGSSDAGTSYIFSGNSVSTGKSQTTGQI 915

Qy    242 AGGAIFSPICSLTGNRGNIVFYNNRC-FKNVETASSEASDGGAIKVT-------TRLDVT 293
Db    916 AGGAIYSPTVTL---NCPATFSNNTASIATPKTSSEDGSSGNSIKDTIGGAIAGTAITLS 972

Qy    294 G---------------------------------------NRGRIFFSDNITKNYGGAI 313
Db    973 GVSRFSGNTADLGAAIGTLANANTPSATSGSQNSITEKITLENGSFIFERN-QANKRGAI 1031

Qy    314 YAPVVTLVDNGPTYFINNIANNKGGAIYI--DGTSNSKISADRHAIIFNENIVTNVTNAN 371
Db   1032 YSPSVSIKGNNIT-FNQNTSTHDGSAIYFTKDATIESLGS-----VLFTGNNVT-ATQAS 1084

Qy    372 GTSTSANPPRRN-------------------AITVASSSGEILLGAGSSQN--------- 403
Db   1085 SATSGQNTNTANYGAAIFGDPGTTQSSQTDAILTLLASSGNITFSNNSLQNNQGDTPASK 1144

Qy    404 -------------------LIFYDPIEVS--------NAGVSVSFNKEADQ---TGSVVF 433
Db   1145 FCSIAGYVKLSLQAAKGKTISFFDCVHTSTKKTGSTQNVYETLDINKEENSNPYTGTIVF 1204

Qy    434 SGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQ-TGGVVSLGNGAVLS 492
Db   1205 ------SSELHEN--KSYIPQNAILHNGTLVLKEKTELHVVSFEQKEGSKLIMEPGAVLS 1256

Qy    493 CYKNGAGNSASNASITLKHIGLNLSSI-LKSGAEI---PLLWVEPTNNSNN--------- 539
Db   1257 ------NQNIANGALAINGLTIDLSSMGTPQAGEIFSPPELRIVATTSSASGGSGVSSSI 1310

Qy    540 -----------YTADTAATFSLSDVK------LSLIDDYGN---SPYESTDLTHALSSQP 579
Db   1311 PTNPKRISAAVPSGSAATTPTMSENKVFLTGDLTLIDPNGNFYQNPMLGSDLDVPLIKLP 1370

Qy    580 MLSISEASDNQLRSDDMDFSGLNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRF 639
Db   1371 ----TNTSDVQVY--DLTLSGDLFPQKGYMGTWTLD--------------SNPQTGKLQ 1409

Qy    640 HRTLLLTWLPAGYVPSPKHRSPLIANTLWG---NMLLATESL----KNSAELTP-SDHPF 691
Db   1410 ARWTFDTYRRWVYIPRDNH---FYANSILGSQNSMIVVKQGLINNMLNNARFDDIAYNNF 1466

Qy    692 WGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFSL--KFSQTYTKLNERYA 749
Db   1467 W---VSGVGTFLAQQGTPLSEEFSYYSRGTSVAIDAKPRQDFILGAAFSKIVGKTKAIKK 1523

Qy    750 KNNVSSK--NYSCQ-----GEMLFSL---QEGFLLTKLV-GLYSYGDHNCHHFYTQGENL 798
Db   1524 MHNYFHKGSEYSYQASVYGGKFLYFLLNKQHGWALPFLIQGVVSYG-HIKHDTTTLYPSI 1582

Qy    799 TSQGTFRSQTMGG----AVFFDLPMKPFGSTHILTAPFLGALGIYSSL--SHFTEVGAYP 852
Db   1583 HERNKGDWEDLGWLADLRISMDLKEPSKDSSKRITV--YGELE-YSSIRQKQFTEIDYDP 1639

Qy    853 RSFSTKTPLINVLVPIG--VKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLAS 908
Db   1640 RHFD-DCAYRNLSLPVGCAVEGAIMNCNILMYN-KLALAYMPSIYRNNPVCKYRVLSS 1695


Search completed: February 4, 2003, 20:59:48
Job time : 46 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2003, 19:08:07 ; Search time 84 Seconds
(without alignments)
1530.796 Million cell updates/sec

Title: US-09-677-752-2
Perfect score: 5086
Sequence: 1 MKKAFFFFLIGNSLSGLARE.....GFYSSSTFCNYLNGEIALRF 965

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5086	100.0	965	23	AAE23453	Chlamydia trachoma
2	5047	99.2	989	20	AAY37242	Chlamydia trachoma
3	4994.5	98.2	964	21	AAB13634	C. trachomatis pmp
4	4994.5	98.2	964	22	AAG83202	Protein encoded by
5	4994.5	98.2	964	23	ABB94173	Chlamydia protein
6	4919.5	96.7	977	21	AAB13640	C. trachomatis pmp
7	4919.5	96.7	977	22	AAG83208	Protein encoded by
8	4919.5	96.7	977	23	ABB94179	Chlamydia protein
9	4884	96.0	956	23	AAE23486	Chlamydia trachoma
10	2605	51.2	500	23	AAE23472	Chlamydia trachoma
11	1123.5	22.1	931	21	AAY92833	C. pneumoniae CPN1
12	1120.5	22.0	938	23	ABB90528	Chlamydia pneumoni
13	1120.5	22.0	940	20	AAY35082	Chlamydia pneumoni
14	1057.5	20.8	969	20	AAY35083	Chlamydia pneumoni
15	1057.5	20.8	1000	21	AAY92832	C. pneumoniae CPN1
16	1052	20.7	934	23	ABB90529	Chlamydia pneumoni
17	989.5	19.5	963	21	AAY95551	Chlamydia pneumoni
18	947.5	18.6	954	21	AAY92830	C. pneumoniae CPN1
19	915	18.0	788	21	AAY92831	C. pneumoniae CPN1
20	727	14.3	822	20	AAY35088	Chlamydia pneumoni
21	651.5	12.8	1013	20	AAY16738	C. trachomatis F s
22	649.5	12.8	918	21	AAY69369	Amino acid sequenc
23	647	12.7	982	21	AAB13633	C. trachomatis pmp
24	647	12.7	982	22	AAG83201	Protein encoded by
25	647	12.7	982	23	ABB94172	Chlamydia protein
26	647	12.7	1006	21	AAB13639	C. trachomatis pmp
27	647	12.7	1006	22	AAG83207	Protein encoded by
28	647	12.7	1006	23	ABB94178	Chlamydia protein
29	646.5	12.7	1013	20	AAY16737	C. trachomatis B s
30	645	12.7	925	21	AAY99843	Chlamydia pneumoni
31	645	12.7	936	21	AAY99842	Chlamydia pneumoni
32	645	12.7	936	23	ABB90602	Chlamydia pneumoni
33	642	12.6	1012	20	AAY16735	C. trachomatis LGV
34	637.5	12.5	918	20	AAW88422	Chlamydia pneumoni
35	636	12.5	926	23	ABB98228	Chlamydia polypept
36	620	12.2	930	20	AAY35052	Chlamydia pneumoni
37	617	12.1	930	23	ABB90548	Chlamydia pneumoni
38	615	12.1	930	21	AAY90240	Chlamydia antigen
39	608	12.0	930	20	AAW88424	Chlamydia pneumoni
40	605.5	11.9	715	22	AAG83273	Chlamydia trachoma
41	605.5	11.9	715	23	ABB94244	Chlamydia trachoma
42	597	11.7	928	20	AAW88417	Chlamydia pneumoni
43	597	11.7	928	21	AAY90236	Chlamydia antigen
44	597	11.7	928	23	ABB90583	Chlamydia pneumoni
45	597	11.7	949	20	AAY35060	Chlamydia pneumoni

ALIGNMENTS

RESULT 1
AAE23453
ID AAE23453 standard; Protein; 965 AA.
XX
AC AAE23453;
XX
DT 27-AUG-2002 (first entry)
XX
DE Chlamydia trachomatis L2 PMPE protein #1.
XX
KW Polymorphic membrane protein; PMPE; PMPI; vaccine; trachoma; PID; LGV;
KW urethritis; epididymitis; conjunctivitis; cervicitis; cervical cancer;
KW pelvic inflammatory disease; lymphogranuloma venereum; tubal occlusion;
KW endometritis; salpingitis; inflammatory heart disease; cardiomyopathy;
KW infertility; autoimmune myocarditis; atherosclerosis; infection;
KW arthritis.
XX
OS Chlamydia trachomatis.
XX
PN WO200228998-A2.
XX
PD 11-APR-2002.
XX
PF 28-SEP-2001; 2001WO-US30345.
XX
PR 02-OCT-2000; 2000US-0677752.
XX
PA (ANTE-) ANTEX BIOLOGICS INC.
XX
PI Jackson WJ;
XX
DR WPI; 2002-426107/45.
DR N-PSDB; AAD37798.

XX
PT Novel purified Chlamydia polymorphic membrane protein E or I, useful
PT for preparing vaccines for preventing or treating diseases associated
PT with Chlamydia infection such as trachoma, and infertility -
XX
PS Claim 1; Fig 5; 160pp; English.
XX
CC The invention relates to Chlamydia sp. polymorphic membrane protein
CC (PMP) E and PMPI and nucleic acid molecules encoding such proteins.
CC PMPI and PMPE polypeptides, nucleic acids and vaccines are useful for
CC preventing, treating or ameliorating trachoma, conjunctivitis, tubal
CC occlusion, cervicitis, cervical cancer, pelvic inflammatory disease
CC (PID), lymphogranuloma venereum (LGV), endometritis, epididymitis,
CC salpingitis, infertility, reactive arthritis, inflammatory heart
CC disease, urethritis, autoimmune myocarditis, cardiomyopathy and
CC atherosclerosis. They are also useful as reagents for clinical or
CC medical diagnosis of Chlamydia infections. The present sequence is
CC Chlamydia trachomatis L2 PMPE protein.
XX
SQ Sequence 965 AA;

Query Match 100.0%; Score 5086; DB 23; Length 965;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 965; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 60

Qy     61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120

Qy    121 RDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVR 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 RDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVR 180

Qy    181 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 240

Qy    241 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 300

Qy    301 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFN 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFN 360

Qy    361 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 420

Qy    421 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 480

Qy    481 GVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNY 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNY 540

Qy    541 TADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSG 600

Qy    601 LNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRS 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 LNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRS 660

Qy    661 PLIANTLWGNMLLATESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 PLIANTLWGNMLLATESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSG 720

Qy    721 YSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGL 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 YSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGL 780

Qy    781 YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYS 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYS 840

Qy    841 SLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPG 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 SLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPG 900

Qy    901 IATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGE 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 IATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGE 960

Qy    961 IALRF 965
          |||||
Db    961 IALRF 965
```

RESULT 2
AAY37242
ID AAY37242 standard; Protein; 989 AA.
XX
AC AAY37242;
XX
DT 07-OCT-1999 (first entry)
XX
DE Chlamydia trachomatis cellular envelope protein.
XX
KW Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;
KW nongonococcal uretritis; epidymitis; cervicitis; salpingitis;
KW bartholinitis; pneumopathy; venereal lymphogranulomatosis.
XX
OS Chlamydia trachomatis.
XX
PN [redacted].
XX
PD 10-JUN-1999.
XX
PF 27-NOV-1998; 98WO-IB01939.
XX
PR 04-NOV-1998; 98US-0107077.
PR 28-NOV-1997; 97FR-0015041.
PR 17-DEC-1997; 97FR-0016034.
XX
PA (GEST) GENSET.
XX
PI Griffais R;
XX
DR WPI; 1999-371125/31.
XX
PT Genome sequence of Chlamydia trachomatis
XX
PS Disclosure; Page 999-1000; 1755pp; English.
XX
CC AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
CC of Chlamydia trachomatis (see AAZ01425). The polypeptides can be used as
CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
CC can also be used to control growth of the microorganism. Chlamydia
CC trachomatis is responsible for a large number of diseases, e.g. eye
CC diseases such as conventional trachoma, nonendemic trachoma,
CC paratrachoma, and inclusion conjunctivitis; genital diseases such as
CC nongonococcal uretritis, epidymitis, cervicitis, salpingitis,
CC perihepatitis, bartholinitis; pneumopathy in breast feeding infants;
CC and venereal lymphogranulomatosis. The polypeptides of the invention
CC may be of use in treating these diseases.
XX
SQ [redacted] Sequence 989 AA;

Query Match 99.2%; Score 5047; DB 20; Length 989;

```
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches  958;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 60
           ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25 MKKAFFFFLIGSSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 84

Qy      61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 144

Qy     121 RDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVR 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     145 RDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVR 204

Qy     181 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     205 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 264

Qy     241 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 300
           ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db     265 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSESSDGGAIKVTTRLDVTGNRGRIF 324

Qy     301 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFN 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     325 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFN 384

Qy     361 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     385 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 444

Qy     421 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     445 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 504

Qy     481 GVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNY 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     505 GVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNY 564

Qy     541 TADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSG 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     565 TADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSG 624

Qy     601 LNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRS 660
           ||||||||||||:||||||||||||||||||||:||||||||||||||||||||||||||||
Db     625 LNVPHYGWQGLWSWGWAKTQDPEPASSATITDPKKANRFHRTLLLTWLPAGYVPSPKHRS 684

Qy     661 PLIANTLWGNMLLATESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     685 PLIANTLWGNMLLATESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSG 744

Qy     721 YSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGL 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db     745 YSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLAKLVGL 804

Qy     781 YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYS 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     805 YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYS 864

Qy     841 SLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     865 SLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQELE 924

Qy     901 IATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGE 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     925 IATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGE 984

Qy     961 IALRF 965
           |||||
Db     985 IALRF 989


RESULT 3
```

```
AAB13634
ID   AAB13634 standard; Protein; 964 AA.
XX
AC   AAB13634;
XX
DT   02-FEB-2001  (first entry)
XX
DE   C. trachomatis pmpE gene protein.
XX
KW   Chlamydial infection; sexually transmitted disease;
KW   pelvic inflammatory disease; PID; tubal obstruction; infertility;
KW   trachoma; blindness; acute respiratory tract infection;
KW   atherosclerosis; coronary heart disease; antibacterial.
XX
OS   Chlamydia trachomatis.
XX
PN   WO200034483-A2.
XX
PD   15-JUN-2000.
XX
PF   08-DEC-1999;   99WO-US29012.
XX
PR   08-DEC-1998;   98US-0208277.
PR   08-APR-1999;   99US-0288594.
PR   01-OCT-1999;   99US-0410568.
PR   22-OCT-1999;   99US-0426571.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Probst P,  Bhatia A,  Skeiky YAW,  Fling SP,  Jen S,  Stromberg EJ;
XX
DR   WPI; 2000-431303/37.
XX
PT   Isolated polypeptide for diagnosis and treatment of Chlamydia infection
PT   comprises immunogenic portion of Chlamydia antigen, which comprises
PT   amino acid sequence encoded by polynucleotide sequence  -
XX
PS   Claim 2; Pages 184-186; 256pp; English.
XX
CC   The present invention relates to new nucleic acid sequences and the
CC   proteins encoded by the nucleic acid sequences. The encoded proteins
CC   comprise an immunogenic portion of a Chlamydia antigen. The encoded
CC   proteins are useful for the serodiagnosis and treatment of Chlamydia
CC   infection. Chlamdiae are intracellular bacterial pathogens that are
CC   responsible for a wide variety of human infections. C. trachomatis
CC   infection is one of the most common sexually transmitted diseases and can
CC   lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
CC   and infertility. Trachoma due to ocular infection with C. trachomatis is
CC   the leading cause of preventable blindness worldwide. C. pneumonia is a
CC   major cause of acute respiratory tract infections in humans and is also
CC   thought to play a role in the pathogenesis of atherosclerosis and
CC   coronary heart disease. The present sequence is a protein isolated in the
CC   present invention.
XX
SQ   Sequence   964 AA;

  Query Match             98.2%;  Score 4994.5;  DB 21;  Length 964;
  Best Local Similarity   98.4%;  Pred. No. 0;
  Matches  950;  Conservative    7;  Mismatches    7;  Indels    1;  Gaps    1;

Qy       1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 60

Qy      61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120

Qy     121 RDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVR 180
           ||||||||||||||| ||  || || :||||||||||||||||||||||||||||||||:
Db     121 RDTIGPVIFENNTCCRLFTWRNPYAA-DKIREGGAIHAQNLYINHNHDVVGFMKNFSYVQ 179

Qy     181 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 240
```

```
Db    180 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 239

Qy    241 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 299

Qy    301 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFN 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFN 359

Qy    361 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 419

Qy    421 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 479

Qy    481 GVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNY 540
          |||||||||||||||| |:|||||||||||||||||||||||||||||||||||||||||||
Db    480 GVVSLGNGAVLSCYKNGTGDSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNY 539

Qy    541 TADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||:|:::|||||||||
Db    540 TADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLQSENIDFSG 599

Qy    601 LNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRS 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 LNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRS 659

Qy    661 PLIANTLWGNMLLATESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSG 720
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    660 PLIANTLWGNMLLATESLKNSAELTPSGHPFWGITGGGLGMMVYQDPRENHPGFHMRSSG 719

Qy    721 YSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGL 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    720 YSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGL 779

Qy    781 YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYS 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    780 YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYS 839

Qy    841 SLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPG 900
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    840 SLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATHRPQAWTVELAYQPVLYRQEPG 899

Qy    901 IATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGE 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    900 IATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGE 959

Qy    961 IALRF 965
          |||||
Db    960 IALRF 964
```

```
RESULT 4
AAG83202
ID   AAG83202 standard; Protein; 964 AA.
XX
AC   AAG83202;
XX
DT   05-SEP-2001  (first entry)
XX
DE   Protein encoded by Chlamydia trachomatis pmpE gene.
XX
KW   Chlamydia; vaccine; infection; fusion protein; antigen;
KW   pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;
KW   acute respiratory tract infection; Cap1; CT529; OMCB;
KW   polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.
XX
OS   Chlamydia trachomatis.
XX
PN   WO200140474-A2.
XX
PD   07-JUN-2001.
XX
PF   04-DEC-2000; 2000WO-US32919.
XX
PR   03-DEC-1999;   99US-0454684.
PR   19-APR-2000; 2000US-0556877.
PR   20-JUN-2000; 2000US-0598419.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Probst P,  Bhatia A,  Skeiky YAW,  Fling SP,  Scholler J;
XX
DR   WPI; 2001-374831/39.
XX
PT   Chlamydia polypeptides and fusion proteins useful for preventing pelvic
PT   inflammatory disease, trachoma, acute respiratory tract infections,
PT   atherosclerosis and heart disease -
XX
PS   Claim 2; Page 191-193; 295pp; English.
XX
CC   The present sequence is provided in a specification relating to
CC   compounds and methods for the treatment and diagnosis of chlamydial
CC   infection. The compounds provided include polypeptides and fusion
CC   proteins comprising immunogenic portions of Chlamydia antigens
CC   and DNA sequences encoding such polypeptides. They are useful for
CC   vaccinating against chlamydial infection, which causes pelvic
CC   inflammatory disease, trachoma, acute respiratory tract infections,
CC   atherosclerosis and heart disease.
XX
SQ   Sequence   964 AA;

  Query Match             98.2%;  Score 4994.5;  DB 22;  Length 964;
  Best Local Similarity   98.4%;  Pred. No. 0;
  Matches  950;  Conservative    7;  Mismatches    7;  Indels    1;  Gaps    1;

Qy      1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 60

Qy     61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120

Qy    121 RDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVR 180
          |||||||||||||||| ||  || || :|||||||||||||||||||||||||||||||:
Db    121 RDTIGPVIFENNTCCRLFTWRNPYAA-DKIREGGAIHAQNLYINHNHDVVGFMKNFSYVQ 179

Qy    181 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 239

Qy    241 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 299

Qy    301 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFN 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFN 359

Qy    361 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 419

Qy    421 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 479

Qy    481 GVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNY 540
          |||||||||||||||| |:|||||||||||||||||||||||||||||||||||||||||||
Db    480 GVVSLGNGAVLSCYKNGTGDSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNY 539
```

```
Qy     541 TADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSG 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||:|:::||||
Db     540 TADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLQSENIDFSG 599

Qy     601 LNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRS 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     600 LNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRS 659

Qy     661 PLIANTLWGNMLLATESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSG 720
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     660 PLIANTLWGNMLLATESLKNSAELTPSGHPFWGITGGGLGMMVYQDPRENHPGFHMRSSG 719

Qy     721 YSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGL 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     720 YSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGL 779

Qy     781 YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYS 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     780 YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYS 839

Qy     841 SLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPG 900
           |||||||||||||||||||||||||||||-|||| ||||||||||||||||||||||||
Db     840 SLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATHRPQAWTVELAYQPVLYRQEPG 899

Qy     901 IATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGE 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     900 IATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGE 959

Qy     961 IALRF 965
           |||||
Db     960 IALRF 964
```

RESULT 5
ABB94173
ID ABB94173 standard; Protein; 964 AA.
XX
AC ABB94173;
XX
DT 05-JUN-2002 (first entry)
XX
DE Chlamydia protein sequence SEQ ID NO:177.
XX
KW Chlamydial infection; Chlamydia; vaccine; detection; diagnosis;
KW antigen; antibacterial; immunostimulant; immune response;
KW Chlamydia-specific T-cell response.
XX
OS Chlamydia sp.
XX
PN WO200208267-A2.
XX
PD 31-JAN-2002.
XX
PF 20-JUL-2001; 2001WO-US23121.
XX
PR 20-JUL-2000; 2000US-0620412.
PR 23-APR-2001; 2001US-0841132.
XX
PA (CORI-) CORIXA CORP.
XX
PI Fling SP, Skeiky YAW, Probst P, Bhatia A;
XX
DR WPI; 2002-179901/23.
XX
PT Novel compositions comprising Chlamydia Cap1 protein and its use in the
PT treatment of Chlamydia infection -
XX
PS Example 1; Page 223-225; 537pp; English.
XX
CC The present invention describes compositions comprising a Chlamydia Cap1
CC protein and methods for the diagnosis and therapy of Chlamydia infection.
CC Chlamydia DNA and protein sequences from the present invention can have
CC antibacterial and immunostimulant activities, and can be used in
CC vaccines. Compounds from the present invention can be used for eliciting
CC an immune response, specifically stimulating a Chlamydia-specific T-cell
CC response or inhibiting the development of a Chlamydia infection in an
CC animal. Methods from the present invention can be used: for detecting the
CC presence of Chlamydia in a patient; to stimulate and/or expand T cells
CC specific for a Chlamydia protein; and for treatment of a Chlamydia
CC infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 964 AA;

```
  Query Match             98.2%;  Score 4994.5;  DB 23;  Length 964;
  Best Local Similarity   98.4%;  Pred. No. 0;
  Matches  950;  Conservative    7;  Mismatches    7;  Indels    1;  Gaps    1;

Qy       1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 60

Qy      61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120

Qy     121 RDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVR 180
           |||||||||||||||  ||  || || :|||||||||||||||||||||||||||||||:
Db     121 RDTIGPVIFENNTCCRLFTWRNPYAA-DKIREGGAIHAQNLYINHNHDVVGFMKNFSYVQ 179

Qy     181 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 239

Qy     241 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     240 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 299

Qy     301 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFN 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     300 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFN 359

Qy     361 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     360 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 419

Qy     421 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     420 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 479

Qy     481 GVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNY 540
           |||||||||||||||| |:|||||||||||||||||||||||||||||||||||||||||
Db     480 GVVSLGNGAVLSCYKNGTGDSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNY 539

Qy     541 TADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSG 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||:|:::||||
Db     540 TADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLQSENIDFSG 599

Qy     601 LNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRS 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     600 LNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRS 659

Qy     661 PLIANTLWGNMLLATESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSG 720
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     660 PLIANTLWGNMLLATESLKNSAELTPSGHPFWGITGGGLGMMVYQDPRENHPGFHMRSSG 719

Qy     721 YSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGL 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     720 YSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGL 779

Qy     781 YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYS 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     780 YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYS 839

Qy     841 SLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPG 900
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
```

```
Db    840 SLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATHRPQAWTVELAYQPVLYRQEPG 899

Qy    901 IATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGE 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    900 IATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGE 959

Qy    961 IALRF 965
          |||||
Db    960 IALRF 964
```

RESULT 6
AAB13640
ID AAB13640 standard; Protein; 977 AA.
XX
AC AAB13640;
XX
DT 02-FEB-2001 (first entry)
XX
DE C. trachomatis pmpE gene amino terminus minus signal sequence protein.
XX
KW Chlamydial infection; sexually transmitted disease;
KW pelvic inflammatory disease; PID; tubal obstruction; infertility;
KW trachoma; blindness; acute respiratory tract infection;
KW atherosclerosis; coronary heart disease; antibacterial.
XX
OS Chlamydia trachomatis.
XX
PN WO200034483-A2.
XX
PD 15-JUN-2000.
XX
PF 08-DEC-1999; 99WO-US29012.
XX
PR 08-DEC-1998; 98US-0208277.
PR 08-APR-1999; 99US-0288594.
PR 01-OCT-1999; 99US-0410568.
PR 22-OCT-1999; 99US-0426571.
XX
PA (CORI-) CORIXA CORP.
XX
PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
XX
DR WPI; 2000-431303/37.
XX
PT Isolated polypeptide for diagnosis and treatment of Chlamydia infection
PT comprises immunogenic portion of Chlamydia antigen, which comprises
PT amino acid sequence encoded by polynucleotide sequence -
XX
PS Claim 2; Pages 210-212; 256pp; English.
XX
CC The present invention relates to new nucleic acid sequences and the
CC proteins encoded by the nucleic acid sequences. The encoded proteins
CC comprise an immunogenic portion of a Chlamydia antigen. The encoded
CC proteins are useful for the serodiagnosis and treatment of Chlamydia
CC infection. Chlamdiae are intracellular bacterial pathogens that are
CC responsible for a wide variety of human infections. C. trachomatis
CC infection is one of the most common sexually transmitted diseases and can
CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
CC and infertility. Trachoma due to ocular infection with C. trachomatis is
CC the leading cause of preventable blindness worldwide. C. pneumonia is a
CC major cause of acute respiratory tract infections in humans and is also
CC thought to play a role in the pathogenesis of atherosclerosis and
CC coronary heart disease. The present sequence is a protein isolated in the
CC present invention.
XX
SQ Sequence 977 AA;

Query Match 96.7%; Score 4919.5; DB 21; Length 977;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 935; Conservative 7; Mismatches 7; Indels 1; Gaps 1;

```
Qy     16 GLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEG 75
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     29 GLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEG 88

Qy     76 AAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCC 135
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     89 AAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCC 148

Qy    136 RPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSEN 195
          | ||  || || :|||||||||||||||||||||||||||||:|||||||||||||||||
Db    149 RLFTWRNPYAA-DKIREGGAIHAQNLYINHNHDVVGFMKNFSYVQGGAISTANTFVVSEN 207

Qy    196 QSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNACCAGGAIFSPICSLTG 255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    208 QSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNACCAGGAIFSPICSLTG 267

Qy    256 NRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYA 315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    268 NRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYA 327

Qy    316 PVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTST 375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    328 PVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTST 387

Qy    376 SANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVSFNKEADQTGSVVFSG 435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    388 SANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVSFNKEADQTGSVVFSG 447

Qy    436 ATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYK 495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    448 ATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYK 507

Qy    496 NGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKL 555
          || |:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    508 NGTGDSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKL 567

Qy    556 SLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSGLNVPHYGWQGLWTWG 615
          ||||||||||||||||||||||||||||||||||:|:::|||||||||||||||||||||
Db    568 SLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLQSENIDFSGLNVPHYGWQGLWTWG 627

Qy    616 WAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLAT 675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    628 WAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLAT 687

Qy    676 ESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFSL 735
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    688 ESLKNSAELTPSGHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFSL 747

Qy    736 KFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQG 795
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    748 KFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQG 807

Qy    796 ENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSF 855
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    808 ENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSF 867

Qy    856 STKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGS 915
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    868 STKTPLINVLVPIGVKGSFMNATHRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGS 927

Qy    916 GSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    928 GSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 977
```

RESULT 7
AAG83208
ID AAG83208 standard; Protein; 977 AA.
XX
AC AAG83208;
XX
DT 05-SEP-2001 (first entry)
XX
DE Protein encoded by Chlamydia trachomatis pmpE gene.

```
XX
KW   Chlamydia; vaccine; infection; fusion protein; antigen;
KW   pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;
KW   acute respiratory tract infection; Capl; CT529; OMCB;
KW   polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.
XX
OS   Chlamydia trachomatis.
XX
PN   WO200140474-A2.
XX
PD   07-JUN-2001.
XX
PF   04-DEC-2000; 2000WO-US32919.
XX
PR   03-DEC-1999;   99US-0454684.
PR   19-APR-2000; 2000US-0556877.
PR   20-JUN-2000; 2000US-0598419.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Probst P,  Bhatia A,  Skeiky YAW,  Fling SP,  Scholler J;
XX
DR   WPI; 2001-374831/39.
XX
PT   Chlamydia polypeptides and fusion proteins useful for preventing pelvic
PT   inflammatory disease, trachoma, acute respiratory tract infections,
PT   atherosclerosis and heart disease -
XX
PS   Claim 2; Page 216-218; 295pp; English.
XX
CC   The present sequence is provided in a specification relating to
CC   compounds and methods for the treatment and diagnosis of chlamydial
CC   infection. The compounds provided include polypeptides and fusion
CC   proteins comprising immunogenic portions of Chlamydia antigens
CC   and DNA sequences encoding such polypeptides. They are useful for
CC   vaccinating against chlamydial infection, which causes pelvic
CC   inflammatory disease, trachoma, acute respiratory tract infections,
CC   atherosclerosis and heart disease.
XX
SQ   Sequence   977 AA;

  Query Match            96.7%;  Score 4919.5;  DB 22;  Length 977;
  Best Local Similarity  98.4%;  Pred. No. 0;
  Matches  935;  Conservative    7;  Mismatches    7;  Indels    1;  Gaps    1;

Qy      16 GLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEG 75
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      29 GLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEG 88

Qy      76 AAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCC 135
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      89 AAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCC 148

Qy     136 RPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSEN 195
           | ||  || || :|||||||||||||||||||||||||||||:|||||||||||||||||
Db     149 RLFTWRNPYAA-DKIREGGAIHAQNLYINHNHDVVGFMKNFSYVQGGAISTANTFVVSEN 207

Qy     196 QSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNACCAGGAIFSPICSLTG 255
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     208 QSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNACCAGGAIFSPICSLTG 267

Qy     256 NRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYA 315
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     268 NRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYA 327

Qy     316 PVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTST 375
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     328 PVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTST 387

Qy     376 SANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVSFNKEADQTGSVVFSG 435
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     388 SANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVSFNKEADQTGSVVFSG 447

Qy     436 ATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYK 495
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     448 ATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYK 507

Qy     496 NGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKL 555
           || |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     508 NGTGDSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKL 567

Qy     556 SLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSGLNVPHYGWQGLWTWG 615
           ||||||||||||||||||||||||||||||||:|:::|||||||||||||||||||||||
Db     568 SLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLQSENIDFSGLNVPHYGWQGLWTWG 627

Qy     616 WAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLAT 675
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     628 WAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLAT 687

Qy     676 ESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFSL 735
           |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db     688 ESLKNSAELTPSGHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFSL 747

Qy     736 KFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQG 795
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     748 KFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQG 807

Qy     796 ENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSF 855
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     808 ENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSF 867

Qy     856 STKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGS 915
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     868 STKTPLINVLVPIGVKGSFMNATHRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGS 927

Qy     916 GSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     928 GSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 977


RESULT 8
ABB94179
ID   ABB94179 standard; Protein; 977 AA.
XX
AC   ABB94179;
XX
DT   05-JUN-2002  (first entry)
XX
DE   Chlamydia protein sequence SEQ ID NO:191.
XX
KW   Chlamydial infection; Chlamydia; vaccine; detection; diagnosis;
KW   antigen; antibacterial; immunostimulant; immune response;
KW   Chlamydia-specific T-cell response.
XX
OS   Chlamydia sp.
XX
PN   WO200208267-A2.
XX
PD   31-JAN-2002.
XX
PF   20-JUL-2001; 2001WO-US23121.
XX
PR   20-JUL-2000; 2000US-0620412.
PR   23-APR-2001; 2001US-0841132.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Fling SP,  Skeiky YAW,  Probst P,  Bhatia A;
XX
DR   WPI; 2002-179901/23.
XX
PT   Novel compositions comprising Chlamydia Capl protein and its use in the
PT   treatment of Chlamydia infection -
XX
PS   Example 1; Page 247-249; 537pp; English.
XX
```

```
CC   The present invention describes compositions comprising a Chlamydia Cap1
CC   protein and methods for the diagnosis and therapy of Chlamydia infection.
CC   Chlamydia DNA and protein sequences from the present invention can have
CC   antibacterial and immunostimulant activities, and can be used in
CC   vaccines. Compounds from the present invention can be used for eliciting
CC   an immune response, specifically stimulating a Chlamydia-specific T-cell
CC   response or inhibiting the development of a Chlamydia infection in an
CC   animal. Methods from the present invention can be used: for detecting the
CC   presence of Chlamydia in a patient; to stimulate and/or expand T cells
CC   specific for a Chlamydia protein; and for treatment of a Chlamydia
CC   infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent
CC   sequences used in the exemplification of the present invention.
XX
SQ   Sequence   977 AA;

  Query Match            96.7%;  Score 4919.5;  DB 23;  Length 977;
  Best Local Similarity  98.4%;  Pred. No. 0;
  Matches  935;  Conservative    7;  Mismatches    7;  Indels    1;  Gaps    1;

Qy      16 GLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEG 75
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      29 GLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEG 88

Qy      76 AAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCC 135
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      89 AAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCC 148

Qy     136 RPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSEN 195
           | ||  || || :|||||||||||||||||||||||||||||:|||||||||||||||||
Db     149 RLFTWRNPYAA-DKIREGGAIHAQNLYINHNHDVVGFMKNFSYVQGGAISTANTFVVSEN 207

Qy     196 QSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNACCAGGAIFSPICSLTG 255
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     208 QSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNACCAGGAIFSPICSLTG 267

Qy     256 NRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYA 315
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     268 NRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYA 327

Qy     316 PVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTST 375
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     328 PVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTST 387

Qy     376 SANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVSFNKEADQTGSVVFSG 435
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     388 SANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVSFNKEADQTGSVVFSG 447

Qy     436 ATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYK 495
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     448 ATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYK 507

Qy     496 NGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKL 555
           || |:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     508 NGTGDSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKL 567

Qy     556 SLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSGLNVPHYGWQGLWTWG 615
           ||||||||||||||||||||||||||||||:|:::|||||||||||||||||||||||||
Db     568 SLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLQSENIDFSGLNVPHYGWQGLWTWG 627

Qy     616 WAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLAT 675
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     628 WAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLAT 687

Qy     676 ESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFSL 735
           |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db     688 ESLKNSAELTPSGHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFSL 747

Qy     736 KFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQG 795
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     748 KFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQG 807

Qy     796 ENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSF 855
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     808 ENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSF 867

Qy     856 STKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGS 915
           |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     868 STKTPLINVLVPIGVKGSFMNATHRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGS 927

Qy     916 GSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     928 GSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 977


RESULT 9
AAE23486
ID   AAE23486 standard; Protein; 956 AA.
XX
AC   AAE23486;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Chlamydia trachomatis L2 PMPE protein #2.
XX
KW   Polymorphic membrane protein; PMPE; PMPI; vaccine; trachoma; PID; LGV;
KW   urethritis; epididymitis; conjunctivitis; cervicitis; cervical cancer;
KW   pelvic inflammatory disease; lymphogranuloma venereum; tubal occlusion;
KW   endometritis; salpingitis; inflammatory heart disease; cardiomyopathy;
KW   infertility; autoimmune myocarditis; atherosclerosis; infection;
KW   arthritis.
XX
OS   Chlamydia trachomatis.
XX
PN   WO200228998-A2.
XX
PD   11-APR-2002.
XX
PF   28-SEP-2001; 2001WO-US30345.
XX
PR   02-OCT-2000; 2000US-0677752.
XX
PA   (ANTE-) ANTEX BIOLOGICS INC.
XX
PI   Jackson WJ;
XX
DR   WPI; 2002-426107/45.
DR   N-PSDB; AAD37835.
XX
PT   Novel purified Chlamydia polymorphic membrane protein E or I, useful
PT   for preparing vaccines for preventing or treating diseases associated
PT   with Chlamydia infection such as trachoma, and infertility -
XX
PS   Claim 1; Fig 8; 160pp; English.
XX
CC   The invention relates to Chlamydia sp. polymorphic membrane protein
CC   (PMP) E and PMPI and nucleic acid molecules encoding such proteins.
CC   PMPI and PMPE polypeptides, nucleic acids and vaccines are useful for
CC   preventing, treating or ameliorating trachoma, conjunctivitis, tubal
CC   occlusion, cervicitis, cervical cancer, pelvic inflammatory disease
CC   (PID), lymphogranuloma venereum (LGV), endometritis, epididymitis,
CC   salpingitis, infertility, reactive arthritis, inflammatory heart
CC   disease, urethritis, autoimmune myocarditis, cardiomyopathy and
CC   atherosclerosis. They are also useful as reagents for clinical or
CC   medical diagnosis of Chlamydia infections. The present sequence is
CC   Chlamydia trachomatis L2 PMPE protein.
XX
SQ   Sequence   956 AA;

  Query Match            96.0%;  Score 4884;  DB 23;  Length 956;
  Best Local Similarity  99.0%;  Pred. No. 0;
  Matches  925;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      32 VPDPTKESLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQK 91
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 VPDPTKESLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQK 82
```

```
Qy      92 EGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNPNAAVNKIR 151
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 EGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNPNAAVNKIR 142

Qy     152 EGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSCFLFMDNICIQTNT 211
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 EGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSCFLFMDNICIQTNT 202

Qy     212 AGKGGAIYAGTSNSFESNNCDLFFINNACCAGGAIFSPICSLTGNRGNIVFYNNRCFKNV 271
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 AGKGGAIYAGTSNSFESNNCDLFFINNACCAGGAIFSPICSLTGNRGNIVFYNNRCFKNV 262

Qy     272 ETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINN 331
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     263 ETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINN 322

Qy     332 IANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSS 391
           :|||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db     323 VANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTSANGTSTSANPPRRNAITVASSS 382

Qy     392 GEILLGAGSSQNLIFYDPIEVSNAGVSVSFNKEADQTGSVVFSGATVNSADFHQRNLQTK 451
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     383 GEILLGAGSSQNLIFYDPIEVSNAGVSVSFNKEADQTGSVVFSGATVNSADFHQRNLQTK 442

Qy     452 TPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNSASNASITLKH 511
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     443 TPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNSASNASITLKH 502

Qy     512 IGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDL 571
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     503 IGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDL 562

Qy     572 THALSSQPMLSISEASDNQLRSDDMDFSGLNVPHYGWQGLWTWGWAKTQDPEPASSATIT 631
           ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db     563 THALSSQPMLSISEASDNQLRSDDMDFSGLNVPHYGWQGLWSWGWAKTQDPEPASSATIT 622

Qy     632 DPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSDHPF 691
           ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     623 DPKKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSDHPF 682

Qy     692 WGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFSLKFSQTYTKLNERYAKN 751
           |||||||||||||:|||||||||||||| ||||||||||||||||||||||||||||||||
Db     683 WGITGGGLGMMVYQEPRENHPGFHMRSSGYFAGMIAGQTHTFSLKFSQTYTKLNERYAKN 742

Qy     752 NVSSKNYSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQGENLTSQGTFRSQTMGG 811
           |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db     743 NVSSKNYSCQGEMLFSLQEGFLLAKLVGLYSYGDHNCHHFYTQGENLTSQGTFRSQTMGG 802

Qy     812 AVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVK 871
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     803 AVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVK 862

Qy     872 GSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQ 931
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     863 GSFMNATQRPQAWTVELAYQPVLYRQELEIATQLLASKGIWFGSGSPSSRHAMSYKISQQ 922

Qy     932 TQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
           |||||||||||||||||||||||||||||||||
Db     923 TQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 956


RESULT 10
AAE23472
ID   AAE23472 standard; Protein; 500 AA.
XX
AC   AAE23472;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Chlamydia trachomatis L2 PMPE protein fragment #18.
XX
KW   Polymorphic membrane protein; PMPE; PMPI; vaccine; trachoma; PID; LGV;
KW   urethritis; epididymitis; conjunctivitis; cervicitis; cervical cancer;
KW   pelvic inflammatory disease; lymphogranuloma venereum; tubal occlusion;
KW   endometritis; salpingitis; inflammatory heart disease; cardiomyopathy;
KW   infertility; autoimmune myocarditis; atherosclerosis; infection;
KW   arthritis.
XX
OS   Chlamydia trachomatis.
XX
PN   WO200228998-A2.
XX
PD   11-APR-2002.
XX
PF   28-SEP-2001; 2001WO-US30345.
XX
PR   02-OCT-2000; 2000US-0677752.
XX
PA   (ANTE-) ANTEX BIOLOGICS INC.
XX
PI   Jackson WJ;
XX
DR   WPI; 2002-426107/45.
DR   N-PSDB; AAD37817.
XX
PT   Novel purified Chlamydia polymorphic membrane protein E or I, useful
PT   for preparing vaccines for preventing or treating diseases associated
PT   with Chlamydia infection such as trachoma, and infertility -
XX
PS   Claim 8; Page 134-136; 160pp; English.
XX
CC   The invention relates to Chlamydia sp. polymorphic membrane protein
CC   (PMP) E and PMPI and nucleic acid molecules encoding such proteins.
CC   PMPI and PMPE polypeptides, nucleic acids and vaccines are useful for
CC   preventing, treating or ameliorating trachoma, conjunctivitis, tubal
CC   occlusion, cervicitis, cervical cancer, pelvic inflammatory disease
CC   (PID), lymphogranuloma venereum (LGV), endometritis, epididymitis,
CC   salpingitis, infertility, reactive arthritis, inflammatory heart
CC   disease, urethritis, autoimmune myocarditis, cardiomyopathy and
CC   atherosclerosis. They are also useful as reagents for clinical or
CC   medical diagnosis of Chlamydia infections. The present sequence is
CC   Chlamydia trachomatis L2 PMPE protein fragment.
XX
SQ   Sequence   500 AA;

  Query Match           51.2%;  Score 2605;  DB 23;  Length 500;
  Best Local Similarity  100.0%;  Pred. No. 3.1e-185;
  Matches  500;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 60

Qy      61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120

Qy     121 RDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVR 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 RDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVR 180

Qy     181 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 240

Qy     241 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 300

Qy     301 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFN 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFN 360

Qy     361 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 420
```

```
Qy    421 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 480

Qy    481 GVVSLGNGAVLSCYKNGAGN 500
          ||||||||||||||||||||
Db    481 GVVSLGNGAVLSCYKNGAGN 500


RESULT 11
AAY92833
ID   AAY92833 standard; Protein; 931 AA.
XX
AC   AAY92833;
XX
DT   29-AUG-2000  (first entry)
XX
DE   C. pneumoniae CPN100628 antigen.
XX
KW   Antigen; anti-inflammatory; respiratory; antibacterial; anti-asthmatic;
KW   anti-arteriosclerotic; vaccine.
XX
OS   Chlamydia pneumoniae.
XX
PN   WO200024765-A2.
XX
PD   04-MAY-2000.
XX
PF   28-OCT-1999;   99WO-CA00992.
XX
PR   28-OCT-1998;   98US-0106034.
PR   28-OCT-1998;   98US-0106039.
PR   28-OCT-1998;   98US-0106042.
PR   28-OCT-1998;   98US-0106044.
PR   29-OCT-1998;   98US-0106072.
PR   29-OCT-1998;   98US-0106073.
PR   29-OCT-1998;   98US-0106074.
PR   29-OCT-1998;   98US-0106087.
PR   02-NOV-1998;   98US-0106587.
PR   02-NOV-1998;   98US-0106588.
PR   02-NOV-1998;   98US-0107034.
PR   02-NOV-1998;   98US-0107035.
XX
PA   (CONN-) CONNAUGHT LAB LTD.
XX
PI   Murdin AD,  Oomen RP,  Wang J;
XX
DR   WPI; 2000-350688/30.
DR   N-PSDB; AAA28712, AAA28713.
XX
PT   Chlamydia antigenes and the proteins they encode, useful for
PT   vaccinating against Chlamydia infections that affect the respiratory
PT   tract
XX
PS   Claim 13; Fig 23; 226pp; English.
XX
CC   The nucleic acids may be used for the recombinant production of the
CC   Chlamydia polypeptides (either in vivo or in vitro) according to standard
CC   recombinant DNA methodologies. The polypeptides may then be used to
CC   vaccinate against Chlamydia infections in mammals. Chlamydia, such as
CC   C. pneumoniae, are pathogens responsible for upper respiratory tract
CC   infections such as community acquired pneumonia, acute respiratory
CC   disease and bronchitis and may be implicated in atherosclerotic changes
CC   and asthma. The nucleic acids may also be used as probes for detecting
CC   the presence of Chlamydia nucleic acids in samples (and therefore
CC   diagnose infections) and the proteins may be used as antigens for the
CC   production of antibodies that may be used to detect Chlamydia proteins
CC   in samples (e.g. via enzyme linked immunosorbant assay (ELISA)).
XX
SQ   Sequence   931 AA;

  Query Match             22.1%;  Score 1123.5;  DB 21;  Length 931;
  Best Local Similarity   30.9%;  Pred. No. 1.5e-74;
  Matches  308;  Conservative  164;  Mismatches  417;  Indels  109;  Gaps   31;

Qy      6 FFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLT---------NC 56
          | |:: |   ||   : : | | |     |            | |||           |
Db      5 FTFVLAN--EGLQLPLETYITLSPEYQAAPQ----------VGFTHNQNQDLAIVGNHND 52

Qy     57 YLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSP 116
          :: : :|         : | |:|  : |         ::| ||: | ||||| ||: |
Db     53 FILDYKYY-------RSNGGALTCKNLL--ISENIGNVFFEKNVCPNSGGAI-YAAQNC- 101

Qy    117 TVEIRDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNF 176
              ||     :|        | || |    :  |||: | |  | :|     |: |
Db    102 ------TISK--NQNYAFTTNLVSDNPTATAGSLL-GGALFAINCSITNNLGQGTFVDNL 152

Qy    177 SYVRGGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFI 236
          :  :|||: |     : :|:   :   | :  |:   || ||:| | : | |:  :
Db    153 ALNKGGALYTETNLSIKDNKGPIIIKQNRAL--NSDSLGGGIYSGNSLNIEGNSGAIQIT 210

Qy    237 NNACCAGGAIFS-PICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGN 295
          :|:  :|| ||| ::: |:  |     |  | |     |  : ||    ||   :  |
Db    211 SNSSGSGGGIFSTQTLTISSNKKLIEISENSAFAN--NYGSNFNPGGGGLTTTFCTILNN 268

Qy    296 RGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYID-GTSNSK--ISA 352
          |  : |::| ::: ||||:|  : : :||| ||:|| |   |  : :  |: |    :||
Db    269 REGVLFNNNQSQSNGGAIHAKSIIIKENGPVYFLNNTATRGGALLNLSAGSGNGSFILSA 328

Qy    353 DRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEV 412
          |   |||| |          :  : ||| |||  :: :  : :||     ::||||||
Db    329 DNGDIIFNNN--------TASKHALNPPYRNAIH-STPNMNLQIGARPGYRVLFYDPIEH 379

Qy    413 S-NAGVSVSFNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQL 471
             :   : || |   ||:|:|||  |:     : |  :       |  | | :|| | |
Db    380 ELPSSFPILFNFETGHTGTVLFSGEHVHQNFTDEMNFFSYLRNTSELRQGVLAVEDGAGL 439

Qy    472 TVNRFTQTGGVVSLGNGAVLSCYKNGAGN--------SASNASITLKHIGLNLSSILKSG 523
             :| | || : || |||::     ||         :   ::||| || ::| |||
Db    440 ACYKFFQRGGTLLLGQGAVITT----AGTIPTPSSTPTTVGSTITLNHIAIDLPSILSFQ 495

Qy    524 AEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSI 583
          |: | :|: ||   : || |:  | ::|   |:| :     ||:| ||:|:|   |:| |
Db    496 AQAPKIWIYPTKTGSTYTEDSNPTITISGT-LTLRNSNNEDPYDSLDLSHSLEKVPLLYI 554

Qy    584 SEASDNQLRSDDMDFSGLNV-PHYGWQGLWTWGWAKTQDPEPASSATITDPQK---ANRF 639
           : :  :: |  :| | ||   |||:||:|:  | :|         |||:|     ||
Db    555 VDVAAQKINSSQLDLSTLNSGEHYGYQGIWSTYWVET--------TTITNPTSLLGANTK 606

Qy    640 HRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSDHPFWGITGG-- 697
          |: |   | | || | |: |   | | || :   |   | : :         |    |
Db    607 HKLLYANWSPLGYRPHPERRGEFITNALWQSAYTALAGLHSLSS--------WDEEKGHA 658

Qy    698 ----GLGMMVYQDPRENHPGFHMRSSGYSAGMIA--GQTHTFSLKFSQTYTKLNERYAKN 751
              |:|::|:|  :    ||    :|||| |  |:  ||| |:| ::|  |  ::|
Db    659 ASLQGIGLLVHQKDKNGFKGFRSHMTGYSATTEATSSQSPNFSLGFAQFFSKAKEHESQN 718

Qy    752 NVSSKNY---SCQGEMLFSLQEGFLLTKLVGL-YSYGDHNCHHFYTQGENLTSQGTFRSQ 807
          : || :|    |  :  :|||    |:  | | | :   : |  |       |||:| :
Db    719 STSSHHYFSGMCIAK--YSLQRVIRLS--VSLAYMFTSEHTHTMYQGLLEGNSQGSFHNH 774

Qy    808 TMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVP 867
          |: ||:      :| | : :   ||: || |  :|: | | | : | ||   || :| :|
Db    775 TLAGALSCVFLPQPHGES-LQIYPFITALAIRGNLAAFQESGDHAREFSLHRPLTDVSLP 833

Qy    868 IGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYK 927
          :|:: |: |  : | |  |::|:  ||||:| : ::|| |:| |    :| : :|:  |
Db    834 VGIRASWKNHHRVPLVWLTEISYRSTLYRQDPELHSKLLISQGTWTTQATPVTYNALGIK 893

Qy    928 ISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
          :   |    :||   |    ||||  :|||    :||
Db    894 VKNTMQVFPKVTLSLDYSADISSSTLSHYLNVASRMRF 931


RESULT 12
```

```
ABB90528
ID   ABB90528 standard; Protein; 938 AA.
XX
AC   ABB90528;
XX
DT   29-JUL-2002  (first entry)
XX
DE   Chlamydia pneumoniae cp6751 protein, SEQ ID NO:5.
XX
KW   Chlamydial infection; antigen; immunogen; vaccine; diagnosis;
KW   human respiratory disease; cardiovascular disease; atherosclerosis;
KW   coronary artery disease; carotid artery stenosis; myocardial infarction;
KW   cerebrovascular disease; aortic aneurysm; claudication; stroke;
KW   strain CWL029.
XX
OS   Chlamydia pneumoniae.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..17
FT                   /label= Signal_peptide
FT   Protein         18..938
FT                   /note= "Mature protein"
XX
PN   WO200202606-A2.
XX
PD   10-JAN-2002.
XX
PF   03-JUL-2001; 2001WO-IB01445.
XX
PR   03-JUL-2000; 2000GB-0016363.
PR   11-JUL-2000; 2000GB-0017047.
PR   21-JUL-2000; 2000GB-0017983.
PR   07-AUG-2000; 2000GB-0019368.
PR   18-AUG-2000; 2000GB-0020440.
PR   14-SEP-2000; 2000GB-0022583.
PR   10-NOV-2000; 2000GB-0027549.
PR   22-DEC-2000; 2000GB-0031706.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Ratti G,  Grandi G;
XX
DR   WPI; 2002-154726/20.
DR   N-PSDB; ABL91186.
XX
PT   Novel Chlamydia pneumoniae protein useful in the manufacture of a
PT   medicament for treatment or prevention of infection due to Chlamydia,
PT   preferably Chlamydia pneumoniae, and for diagnostic purposes  -
XX
PS   Claim 1; Page 43; 364pp; English.
XX
CC   Sequences ABB90526-ABB90715 represent novel proteins from Chlamydia
CC   pneumoniae (strain CWL029), and ABL91184-ABL91373 represent DNA encoding
CC   them. The proteins are predicted to be immunogenic and may therefore be
CC   useful in vaccine production and for diagnostic purposes. Chlamydia
CC   pneumoniae is a common cause of respiratory disease in humans, and is
CC   also involved in the development of cardiovascular diseases such as
CC   atherosclerosis, coronary artery disease, carotid artery stenosis,
CC   myocardial infarction, cerebrovascular disease, aortic aneurysm,
CC   claudication and stroke. The proteins and nucleic acids of the invention
CC   may be used in vaccines and pharmaceutical compositions for the
CC   prevention or treatment of chlamydial infections, particularly Chlamydia
CC   pneumoniae infections. The proteins may also be used in the detection of
CC   Chlamydia pneumoniae, and the nucleic acids may be used in PCR, branched
CC   DNA probe assay or blotting techniques for determining Chlamydia
CC   pneumoniae gene expression. The present sequence represents a
CC   specifically claimed Chlamydia pneumoniae protein of the invention.
XX
SQ   Sequence   938 AA;

  Query Match             22.0%;  Score 1120.5;  DB 23;  Length 938;
  Best Local Similarity   30.9%;  Pred. No. 2.5e-74;
  Matches  308;  Conservative  163;  Mismatches  418;  Indels  109;  Gaps   31;

Qy       6 FFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLT---------NC 56
           | |:: |   ||   : : | | |      |                | |||          |
Db      12 FTFVLAN--EGLQLPLETYITLSPEYQAAPQ----------VGFTHNQNQDLAIVGNHND 59

Qy      57 YLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSP 116
           :: : :|         : | |:|  : |         ::| ||: | ||||| ||: |
Db      60 FILDYKYY-------RSNGGALTCKNLL--ISENIGNVFFEKNVCPNSGGAI-YAAQNC- 108

Qy     117 TVEIRDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNF 176
                ||     :|        | || |    :  |||: | |  | :|     |: |
Db     109 ------TISK--NQNYAFTTNLVSDNPTATAGSLL-GGALFAINCSITNNLGQGTFVDNL 159

Qy     177 SYVRGGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFI 236
           :  :|||: |     : :|:   :   |  :  |:   || ||:| | : | |:  :
Db     160 ALNKGGALYTETNLSIKDNKGPIIIKQNRAL--NSDSLGGGIYSGNSLNIEGNSGAIQIT 217

Qy     237 NNACCAGGAIFS-PICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGN 295
           :|:  :|| |||    ::: |:  |    |  | |     |  : ||    ||   :  |
Db     218 SNSSGSGGGIFSTQTLTISSNKKLIEISENSAFAN--NYGSNFNPGGGGLTTTFCTILNN 275

Qy     296 RGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYID-GTSNSK--ISA 352
           |  : |::| ::: ||||:|  : : :||| ||:|| |   |  : :  |: |    :||
Db     276 REGVLFNNNQSQSNGGAIHAKSIIIKENGPVYFLNNTATRGGALLNLSAGSGNGSFILSA 335

Qy     353 DRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEV 412
           |   |||| |          :  : ||| ||||  :: :  : :||     ::||||||
Db     336 DNGDIIFNNN--------TASKHALNPPYRNAIH-STPNMNLQIGARPGYRVLFYDPIEH 386

Qy     413 S-NAGVSVSFNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQL 471
              :  : || |   ||:|:|||  |:     : |  :       |  | | :|| | |
Db     387 ELPSSFPILFNFETGHTGTVLFSGEHVHQNFTDEMNFFSYLRNTSELRQGVLAVEDGAGL 446

Qy     472 TVNRFTQTGGVVSLGNGAVLSCYKNGAGN--------SASNASITLKHIGLNLSSILKSG 523
           :| | || : || |||::     ||          :   ::||| || ::| |||
Db     447 ACYKFFQRGGTLLLGQGAVITT----AGTIPTPSSTPTTVGSTITLNHIAIDLPSILSFQ 502

Qy     524 AEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSI 583
           |: | :|: ||   : || |:  | ::|   |:| :     ||:| ||:|:|   |:| |
Db     503 AQAPKIWIYPTKTGSTYTEDSNPTITISGT-LTLRNSNNEDPYDSLDLSHSLEKVPLLYI 561

Qy     584 SEASDNQLRSDDMDFSGLNV-PHYGWQGLWTWGWAKTQDPEPASSATITDPQK---ANRF 639
            : :  :: | :| | ||   |||:||:|:  | :|         {||:|     ||
Db     562 VDVAAQKINSSQLDLSTLNSGEHYGYQGIWSTYWVET--------TTITNPTSLLGANTK 613

Qy     640 HRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSDHPFWGITGG-- 697
           |: |   | | || | |: |   | | || :   |   | : :         |    |
Db     614 HKLLYANWSPLGYRPHPERRGEFITNALWQSAYTALAGLHSLSS--------WDEEKGHA 665

Qy     698 ----GLGMMVYQDPRENHPGFHMRSSGYSAGMIA--GQTHTFSLKFSQTYTKLNERYAKN 751
               |:|::|:|  :    ||    :|||   |   |:  ||| |:| ::|  |  ::|
Db     666 ASLQGIGLLVHQKDKNGFKGFRSHMTGYSATTEATSSQSPNFSLGFAQFFSKAKEHESQN 725

Qy     752 NVSSKNY---SCQGEMLFSLQEGFLLTKLVGL-YSYGDHNCHHFYTQGENLTSQGTFRSQ 807
           : || :|    |    ||  :|   |:  | | | :   : |  |       |||:| :
Db     726 STSSHHYFSGMCIENTLF--KEWIRLS--VSLAYMFTSEHTHTMYQGLLEGNSQGSFHNH 781

Qy     808 TMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVP 867
           |: ||:      :| | : :   ||: || |  :|: | | | : | ||   || :| :|
Db     782 TLAGALSCVFLPQPHGES-LQIYPFITALAIRGNLAAFQESGDHAREFSLHRPLTDVSLP 840

Qy     868 IGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYK 927
           :|:: |: |  : |  |  |::|:  ||||:| : ::|| |:| |    :| : :|:  |
Db     841 VGIRASWKNHHRVPLVWLTEISYRSTLYRQDPELHSKLLISQGTWTTQATPVTYNALGIK 900

Qy     928 ISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
           :    |    :||   |    ||||  :|||    :||
Db     901 VKNTMQVFPKVTLSLDYSADISSSTLSHYLNVASRMRF 938


RESULT 13
AAY35082
ID   AAY35082 standard; Protein; 940 AA.
```

```
XX
AC   AAY35082;
XX
DT   13-SEP-1999  (first entry)
XX
DE   Chlamydia pneumoniae surface exposed polypeptide.
XX
KW   Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW   sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW   vaccine; neutralising epitope.
XX
OS   Chlamydia pneumoniae.
XX
PN   WO9927105-A2.
XX
PD   03-JUN-1999.
XX
PF   20-NOV-1998;   98WO-IB01890.
XX
PR   04-NOV-1998;   98US-0107078.
PR   21-NOV-1997;   97FR-0014673.
XX
PA   (GEST ) GENSET.
XX
PI   Griffais R;
XX
DR   WPI; 1999-357842/30.
XX
PT   Genome sequence of Chlamydia pneumoniae
XX
PS   Page 965-967; Disclosure; 1912pp; English.
XX
CC   AAY34584-Y35879 represent the proteins encoded by all the open reading
CC   frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.
CC   C. pneumoniae causes respiratory disease such as pneumonia and
CC   bronchitis and is thought to be a contributing factor in heart
CC   disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC   nodosum or pharyngitis. The polypeptides encoded by the open reading
CC   frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in
CC   immunogenic compositions as vaccines. Vectors containing C. pneumoniae
CC   nucleotides sequences can also be used as immunogenic compositions,
CC   especially where the vector directs the expression of a neutralising
CC   epitope of C. pneumoniae.
XX
SQ   Sequence   940 AA;

  Query Match             22.0%;  Score 1120.5;  DB 20;  Length 940;
  Best Local Similarity   30.9%;  Pred. No. 2.5e-74;
  Matches  308;  Conservative  163;  Mismatches  418;  Indels  109;  Gaps   31;

Qy      6 FFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLT---------NC 56
          | |:: |   ||   : : | | |      |           | |||          |
Db     14 FTFVLAN--EGLQLPLETYITLSPEYQAAPQ----------VGFTHNQNQDLAIVGNHND 61

Qy     57 YLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSP 116
          :: : :|         : | |:|  : |          ::| ||: | ||||| ||: |
Db     62 FILDYKYY-------RSNGGALTCKNLL--ISENIGNVFFEKNVCPNSGGAI-YAAQNC- 110

Qy    117 TVEIRDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNF 176
                ||     :|        | || |    :  |||: | |  | :|     |: |
Db    111 ------TISK--NQNYAFTTNLVSDNPTATAGSLL-GGALFAINCSITNNLGQGTFVDNL 161

Qy    177 SYVRGGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFI 236
          :  :|||: |     : :|:   :    |  :   |:   || ||:| | : | |:  :
Db    162 ALNKGGALYTETNLSIKDNKGPIIIKQNRAL--NSDSLGGGIYSGNSLNIEGNSGAIQIT 219

Qy    237 NNACCAGGAIFS-PICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGN 295
          :|:  :|| |||    ::: |:  |     |  | |      |  : ||    ||   :  |
Db    220 SNSSGSGGGIFSTQTLTISSNKKLIEISENSAFAN--NYGSNFNPGGGGLTTTFCTILNN 277

Qy    296 RGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYID-GTSNSK--ISA 352
          |  : |::| ::: ||||:|  : : :||| ||:|| |   |  : :  |: |    :||
Db    278 REGVLFNNNQSQSNGGAIHAKSIIIKENGPVYFLNNTATRGGALLNLSAGSGNGSFILSA 337

Qy    353 DRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEV 412
          |   |||| |          :  : ||| |||  :: :  : :||     ::||||||
Db    338 DNGDIIFNNN--------TASKHALNPPYRNAIH-STPNMNLQIGARPGYRVLFYDPIEH 388

Qy    413 S-NAGVSVSFNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQL 471
            :   : || |   ||:|:|||  |:      : |  :       |  | | :|| | |
Db    389 ELPSSFPILFNFETGHTGTVLFSGEHVHQNFTDEMNFFSYLRNTSELRQGVLAVEDGAGL 448

Qy    472 TVNRFTQTGGVVSLGNGAVLSCYKNGAGN--------SASNASITLKHIGLNLSSILKSG 523
             :| | || : || |||::     ||         :   ::||| || ::| |||
Db    449 ACYKFFQRGGTLLLGQGAVITT----AGTIPTPSSTPTTVGSTITLNHIAIDLPSILSFQ 504

Qy    524 AEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSI 583
          |: | :|: ||   : || |:  | ::|   |:| :     ||:| ||:|:|   |:| |
Db    505 AQAPKIWIYPTKTGSTYTEDSNPTITISGT-LTLRNSNNEDPYDSLDLSHSLEKVPLLYI 563

Qy    584 SEASDNQLRSDDMDFSGLNV-PHYGWQGLWTWGWAKTQDPEPASSATITDPQK---ANRF 639
            : :  :: |  :| | ||   |||:||:|:  | :|        |||:|     ||
Db    564 VDVAAQKINSSQLDLSTLNSGEHYGYQGIWSTYWVET--------TTITNPTSLLGANTK 615

Qy    640 HRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSDHPFWGITGG-- 697
          |: |   | | || | |: |   | | || :   |   | : :          |    |
Db    616 HKLLYANWSPLGYRPHPERRGEFITNALWQSAYTALAGLHSLSS--------WDEEKGHA 667

Qy    698 ----GLGMMVYQDPRENHPGFHMRSSGYSAGMIA--GQTHTFSLKFSQTYTKLNERYAKN 751
              |:|::|:|  :    ||    :|||   |   |:  ||| |:| ::|  |  ::|
Db    668 ASLQGIGLLVHQKDKNGFKGFRSHMTGYSATTEATSSQSPNFSLGFAQFFSKAKEHESQN 727

Qy    752 NVSSKNY---SCQGEMLFSLQEGFLLTKLVGL-YSYGDHNCHHFYTQGENLTSQGTFRSQ 807
          : || :|    |    ||  :|   |:  | | | :   : |  |       |||:| :
Db    728 STSSHHYFSGMCIENTLF--KEWIRLS--VSLAYMFTSEHTHTMYQGLLEGNSQGSFHNH 783

Qy    808 TMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVP 867
          |: ||:      :| | : :   ||: || |  :|: | | | : | ||   || :| :|
Db    784 TLAGALSCVFLPQPHGES-LQIYPFITALAIRGNLAAFQESGDHAREFSLHRPLTDVSLP 842

Qy    868 IGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYK 927
          :|:: |: |  : |  |  |::|:  ||||:| : ::|| |:| |    :| : :|:  |
Db    843 VGIRASWKNHHRVPLVWLTEISYRSTLYRQDPELHSKLLISQGTWTTQATPVTYNALGIK 902

Qy    928 ISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
          :    |   :||   |    ||||  :|||    :||
Db    903 VKNTMQVFPKVTLSLDYSADISSSTLSHYLNVASRMRF 940


RESULT 14
AAY35083
ID   AAY35083 standard; Protein; 969 AA.
XX
AC   AAY35083;
XX
DT   13-SEP-1999  (first entry)
XX
DE   Chlamydia pneumoniae surface exposed polypeptide.
XX
KW   Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW   sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW   vaccine; neutralising epitope.
XX
OS   Chlamydia pneumoniae.
XX
PN   WO9927105-A2.
XX
PD   03-JUN-1999.
XX
PF   20-NOV-1998;   98WO-IB01890.
XX
PR   04-NOV-1998;   98US-0107078.
PR   21-NOV-1997;   97FR-0014673.
XX
PA   (GEST ) GENSET.
XX
```

```
PI   Griffais R;
XX
DR   WPI; 1999-357842/30.
XX
PT   Genome sequence of Chlamydia pneumoniae
XX
PS   Page 967-969; Disclosure; 1912pp; English.
XX
CC   AAY34584-Y35879 represent the proteins encoded by all the open reading
CC   frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.
CC   C. pneumoniae causes respiratory disease such as pneumonia and
CC   bronchitis and is thought to be a contributing factor in heart
CC   disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC   nodosum or pharyngitis. The polypeptides encoded by the open reading
CC   frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in
CC   immunogenic compositions as vaccines. Vectors containing C. pneumoniae
CC   nucleotides sequences can also be used as immunogenic compositions,
CC   especially where the vector directs the expression of a neutralising
CC   epitope of C. pneumoniae.
XX
SQ   Sequence   969 AA;

  Query Match            20.8%;  Score 1057.5;  DB 20;  Length 969;
  Best Local Similarity  30.5%;  Pred. No. 1.3e-69;
  Matches  302;  Conservative  165;  Mismatches  447;  Indels   77;  Gaps   28;

Qy       6 FFFLIGNSLSGLAREVPSRIFLMPNSVPD-------------PTKESLSNK------ISL 46
Db      25 FLFYILGNFTACMFGMTPAVYSLQTDSLEKFALERDEEFRTSFPLLDSLSTLTGFSPITTF 84

Qy      47 TGDTHNLT-NCYLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESG 105
Db      85 VGNRHNSSQDIVLSNYKSIDNILLLWTSAGGAVSCNNFL--LSNVEDHAFFSKNLAIGTG 142

Qy     106 GAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQ-NLYIN 164
Db     143 GAIACQG----ACTITKNRGPLIFFSNRGLN-------NASTGGETRGGAIACNGDFTIS 191

Qy     165 HNHDVVGFMKNFSYVRGGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSN 224
Db     192 QNQGTFYFVNNSVNNWGGALSTNGHCRIQSNRAPLLFFN------NTAPSGGGALRSENT 245

Qy     225 SFESNNCDLFFINNACCAGGAIFSPI-CSLTGNRGNIVFYNNRCFKNVETASSEASDGGA 283
Db     246 TISDNTRPIYFKNNCGNNGGAIQTSVTVAIKNNSGSVIFNNNTALSG--SINSGNGSGGA 303

Qy     284 IKVTTRLDVTGNRGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYID 343
Db     304 I-YTTNLSIDDNPGTILFNNNYCIRDGGAICTQFLTIKNSGHVYFTNN-QGNWGGALMLL 361

Qy     344 GTSNSKISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQN 403
Db     362 QDSTCLLFAEQGNIAFQNNEVFLTTFG----------RYNAIHCTPNS-NLQLGANKGYT 410

Qy     404 LIFYDPIEVSNAGVS-VSFNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGF 462
Db     411 TAFFDPIEHQHPTTNPLIFNPNANHQGTILFSSAYIPEASDYENNFISSSKNTSELRNGV 470

Qy     463 LCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNSAS-NASITLKHIGLNLSSILK 521
Db     471 LSIEDRAGWQFYKFTQKGGILKLGHAASIATTANSETPSTSVGSQVIINNLAINLPSILA 530

Qy     522 SGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPML 581
Db     531 KG-KAPTLWIRPLQSSAPFTEDNNPTITLSG-PLTLLNEENRDPYDSIDLSEPLQNIHLL 588

Qy     582 SISEASDNQLRSDDMDFSGLN-VPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFH 640
Db     589 SLSDVTARHINTDNFHPESLNATEHYGYQGIWSPYWVET--ITTTNNASI---ETANTLY 643

Qy     641 RTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSD--HPFWGITGGG 698
Db     644 RALYANWTPLGYKVNPEYQGDLATTPLWQSFHTMFSLLRSYNRTGDSDIERPFLEIQGIA 703

Qy     699 LGMMVYQDPRENHPGFHMRSSGYS--AGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSK 756
Db     704 DGLFVHQNSIPGAPGFRIQSTGYSLQASSETSLHQKISLGFAQFFTRTKEIGSSNNVSAH 763

Qy     757 N--YSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVF 814
Db     764 NTVSSLYVELPW-FQEAF-ATSTVLAYGYGDHHLHSLHPSHQE-QAEGTCYSHTLAAAIG 820

Qy     815 FDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSF 874
Db     821 CSFPWQQKSYLHL--SPFVQAIAIRSHQTAFEEIGDNPRKFVSQKPFYNLTLPLGIQGKW 878

Qy     875 MNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQP 934
Db     879 QSKFHVPTEWTLELSYEPVLYQQNPQIGVTLLASGGSWDILGHNYVRNALGYKVHNQTAL 938

Qy     935 LSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
Db     939 FRSLDLFLDYQGSVSSSTSTHHLQAGSTLKF 969


RESULT 15
AAY92832
ID   AAY92832 standard; Protein; 1000 AA.
XX
AC   AAY92832;
XX
DT   07-SEP-2000  (first entry)
XX
DE   C. pneumoniae CPN100626 full-length antigen.
XX
KW   Antigen; anti-inflammatory; respiratory; antibacterial; anti-asthmatic;
KW   anti-arteriosclerotic; vaccine.
XX
OS   Chlamydia pneumoniae.
XX
PN   WO200024765-A2.
XX
PD   04-MAY-2000.
XX
PF   28-OCT-1999;   99WO-CA00992.
XX
PR   28-OCT-1998;   98US-0106034.
PR   28-OCT-1998;   98US-0106039.
PR   28-OCT-1998;   98US-0106042.
PR   28-OCT-1998;   98US-0106044.
PR   29-OCT-1998;   98US-0106072.
PR   29-OCT-1998;   98US-0106073.
PR   29-OCT-1998;   98US-0106074.
PR   29-OCT-1998;   98US-0106087.
PR   02-NOV-1998;   98US-0106587.
PR   02-NOV-1998;   98US-0106588.
PR   02-NOV-1998;   98US-0107034.
PR   02-NOV-1998;   98US-0107035.
XX
PA   (CONN-) CONNAUGHT LAB LTD.
XX
PI   Murdin AD,  Oomen RP,  Wang J;
XX
DR   WPI; 2000-350688/30.
DR   N-PSDB; AAA28710.
XX
PT   Chlamydia antigenes and the proteins they encode, useful for
PT   vaccinating against Chlamydia infections that affect the respiratory
PT   tract
XX
PS   Claim 13; Fig 21; 226pp; English.
XX
CC   The nucleic acids may be used for the recombinant production of the
CC   Chlamydia polypeptides (either in vivo or in vitro) according to standard
CC   recombinant DNA methodologies. The polypeptides may then be used to
CC   vaccinate against Chlamydia infections in mammals. Chlamydia, such as
```

CC C. pneumoniae, are pathogens responsible for upper respiratory tract
CC infections such as community acquired pneumonia, acute respiratory
CC disease and bronchitis and may be implicated in atherosclerotic changes
CC and asthma. The nucleic acids may also be used as probes for detecting
CC the presence of Chlamydia nucleic acids in samples (and therefore
CC diagnose infections) and the proteins may be used as antigens for the
CC production of antibodies that may be used to detect Chlamydia proteins
CC in samples (e.g. via enzyme linked immunosorbant assay (ELISA)).
XX
SQ Sequence 1000 AA;

Query Match 20.8%; Score 1057.5; DB 21; Length 1000;
Best Local Similarity 30.5%; Pred. No. 1.3e-69;
Matches 302; Conservative 166; Mismatches 446; Indels 77; Gaps 28;

```
Qy      6 FFFLIGNSLSGLAREVPSRIFLMPNSVPD----------PTKESLSNK------ISL 46
          | | :||  : :    |:  |  :|:              |  :|||        :
Db     56 FLFYLGNFTACMFGMTPAVYSLQTDSLEKFALERDEEFRTSFPLLDSLSTLTGFSPITTF 115

Qy     47 TGDTHNLT-NCYLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESG 105
           |: || : :  | | : |  ||    : | ||:  ::|      ::  :|:|||   :|
Db    116 VGNRHNSSQDIVLSNYKSIDNILLLWTSAGGAVSCNNFL--LSNVEDHAFFSKNLAIGTG 173

Qy    106 GAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQ-NLYIN 164
          |||          |    ||:|| :|          ||:      ||||    :  |:
Db    174 GAIACQG----ACTITKNRGPLIFFSNRGLN-------NASTGGETRGGAIACNGDFTIS 222

Qy    165 HNHDVVGFMKNFSYVRGGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSN 224
          |     |: |     |||:||     :  |::  || :      |||  ||      :
Db    223 QNQGTFYFVNNSVNNWGGALSTNGHCRIQSNRAPLLFFN------NTAPSGGGALRSENT 276

Qy    225 SFESNNCDLFFINNACCAGGAIFSPI-CSLTGNRGNIVFYNNRCFKNVETASSEASDGGA 283
          :   |   ::| ||    |||| : :  ::  | |:::| ||        : :|    |||
Db    277 TISDNTRPIYFKNNCGNNGGAIQTSVTVAIKNNSGSVIFNNNTALSG--SINSGNGSGGA 334

Qy    284 IKVTTRLDVTGNRGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYID 343
          |  || | :  | | | |::|     ||||    :|: ::|  || ||   | |||: :
Db    335 I-YTTNLSIDDNPGTILFNNNYCIRDGGAICTQFLTIKNSGHVYFTNN-QGNWGGALMLL 392

Qy    344 GTSNSKISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQN 403
            |   : |::  | |  | |   |            | |||    :|  : |||
Db    393 QDSTCLLFAEQGNIAFQNNEVFLTTFG----------RYNAIHCTPNS-NLQLGANKGYT 441

Qy    404 LIFYDPIEVSNAGVS-VSFNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGF 462
          |:||||  :   : : ||  |:  |:::|| | :  |  :: |  : :     | ||
Db    442 TAFFDPIEHQHPTTNPLIFNPNANHQGTILFSSAYIPEASDYENNFISSSKNTSELRNGV 501

Qy    463 LCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNSAS-NASITLKHIGLNLSSILK 521
          | ||| |     :||| ||:: ||: | ::   |    | |  : : : :: :|| |||
Db    502 LSIEDRAGWQFYKFTQKGGILKLGHAASIATTANSETPSTSVGSQVIINNLAINLPSILA 561

Qy    522 SGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPML 581
          | : | ||: |  :|  :| |   | :||   |:|:::    ||:| ||:  | :  :|
Db    562 KG-KAPTLWIRPLQSSAPFTEDNNPTITLSG-PLTLLNEENRDPYDSIDLSEPLQNIHLL 619

Qy    582 SISEASDNQLRSDDMDFSGLN-VPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFH 640
          |:|: :   : :|:     ||   |||:||:|:  | :|      ::|:|   : ||  :
Db    620 SLSDVTARHINTDNFHPESLNATEHYGYQGIWSPYWVET--ITTTNNASI---ETANTLY 674

Qy    641 RTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSD--HPFWGITGGG 698
          | |   | | ||  :|:::  |    || :       |::      ||   ||  | |
Db    675 RALYANWTPLGYKVNPEYQGDLATTPLWQSFHTMFSLLRSYNRTGDSDIERPFLEIQGIA 734

Qy    699 LGMMVYQDPRENHPGFHMRSSGYS--AGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSK 756
          |: |:|:     ||| ::|:||| |          || |:| :|:  |  : ||||:
Db    735 DGLFVHQNSIPGAPGFRIQSTGYSLQASSETSLHQKISLGFAQFFTRTKEIGSSNNVSAH 794

Qy    757 N--YSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVF 814
          |   |   |: :  || |  :  :  | ||||: | :    :|  ::||  | |:  |:
Db    795 NTVSSLYVELPW-FQEAFATSHSLA-YGYGDHHLHAYIRHIKN-RAEGTCYSHTLAAAIG 851

Qy    815 FDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSF 874
            | :     |:  :||: |: | |  : | |:|  || | :: |  |: :|:|::| :
```

```
Db    852 CSFPWQQKSYLHL--SPFVQAIAIRSHQTAFEEIGDNPRKFVSQKPFYNLTLPLGIQGKW 909

Qy    875 MNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQP 934
            :    |  ||:||:|||||:| | |   |||| | |    |    |:|: ||:  ||
Db    910 QSKFHVPTEWTLELSYQPVLYQQNPQIGVTLLASGGSWDILGHNYVRNALGYKVHNQTAL 969

Qy    935 LSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
            |  |   | |  ||||  ::|     |:|
Db    970 FRSLDLFLDYQGSVSSSTSTHHLQAGSTLKF 1000
```

Search completed: February 4, 2003, 20:55:04
Job time : 88 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 4, 2003, 15:51:52 ; Search time 543 Seconds
(without alignments)
12018.952 Million cell updates/sec

Title: US-09-677-752-1
Perfect score: 2898
Sequence: 1 atgaaaaaagcgtttttctt.....aaattgctctgcgattctag 2898

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2898	100.0	2898	24	AAD37798	Chlamydia trachoma
2	2866	98.9	1038602	20	AAZ01425	Complete genome se
3	2821.2	97.3	2895	22	AAH56256	Chlamydia trachoma
4	2821.2	97.3	2895	24	ABL92485	Chlamydia DNA sequ
5	2819.6	97.3	2895	21	AAA64753	C. trachomatis pmp
6	2777.8	95.9	2934	21	AAA64759	C. trachomatis pmp
7	2777.8	95.9	2934	22	AAH56262	Chlamydia trachoma
8	2777.8	95.9	2934	24	ABL92491	Chlamydia DNA sequ
9	2776.8	95.8	2871	24	AAD37835	Chlamydia trachoma
10	1500	51.8	1500	24	AAD37817	Chlamydia trachoma
c 11	1190.6	41.1	1634	22	AAS57011	C. trachomatis DNA
12	321	11.1	321	24	AAD37801	Chlamydia trachoma
13	276	9.5	276	24	AAD37805	Chlamydia trachoma
14	243	8.4	243	24	AAD37802	Chlamydia trachoma
15	228	7.9	228	24	AAD37811	Chlamydia trachoma
16	228	7.9	228	24	AAD37812	Chlamydia trachoma
17	201	6.9	201	24	AAD37804	Chlamydia trachoma
18	198	6.8	198	24	AAD37803	Chlamydia trachoma
19	198	6.8	198	24	AAD37806	Chlamydia trachoma
20	198	6.8	198	24	AAD37808	Chlamydia trachoma
21	183	6.3	183	24	AAD37814	Chlamydia trachoma
22	153	5.3	153	24	AAD37807	Chlamydia trachoma
23	153	5.3	153	24	AAD37815	Chlamydia trachoma
24	136.4	4.7	2805	24	ABL91187	Chlamydia pneumoni
c 25	136.4	4.7	273254	21	AAC81914	Chlamydia pneumoni
26	134.8	4.7	1230025	20	AAX91990	Nucleotide sequenc
27	131.6	4.5	3003	21	AAA28711	C. pneumoniae CPN1
28	131.6	4.5	3200	21	AAA28710	C. pneumoniae CPN1
29	126	4.3	126	24	AAD37800	Chlamydia trachoma
30	115.8	4.0	3092	21	AAA50038	DNA encoding Chlam
31	108	3.7	108	24	AAD37809	Chlamydia trachoma
32	108	3.7	108	24	AAD37810	Chlamydia trachoma
33	93	3.2	93	24	AAD37816	Chlamydia trachoma
34	89.6	3.1	2796	21	AAA28713	C. pneumoniae CPN1
35	89.6	3.1	3000	21	AAA28712	C. pneumoniae CPN1
36	88	3.0	2817	24	ABL91186	Chlamydia pneumoni
37	71.6	2.5	2865	21	AAA28709	C. pneumoniae CPN1
38	71.6	2.5	3150	21	AAA28708	C. pneumoniae CPN1
39	63	2.2	63	24	AAD37813	Chlamydia trachoma
40	48.4	1.7	1511	20	AAX60546	C. trachomatis HMW
41	48.4	1.7	1515	20	AAX60555	Seq ID No: 22 of W
42	48.4	1.7	2949	21	AAA64752	C. trachomatis pmp
43	48.4	1.7	2949	22	AAH56255	Chlamydia trachoma
44	48.4	1.7	2949	24	ABL92484	Chlamydia DNA sequ
45	48.4	1.7	3021	21	AAA64758	C. trachomatis pmp

ALIGNMENTS

RESULT 1
AAD37798
ID AAD37798 standard; DNA; 2898 BP.
XX
AC AAD37798;
XX
DT 27-AUG-2002 (first entry)
XX
DE Chlamydia trachomatis L2 PMPE DNA #1.
XX
KW Polymorphic membrane protein; PMPE; PMPI; vaccine; trachoma; PID; LGV;
KW urethritis; epididymitis; conjunctivitis; cervicitis; cervical cancer;
KW pelvic inflammatory disease; lymphogranuloma venereum; tubal occlusion;
KW endometritis; salpingitis; inflammatory heart disease; cardiomyopathy;
KW infertility; autoimmune myocarditis; atherosclerosis; infection;
KW arthritis; gene; ds.
XX
OS Chlamydia trachomatis.
XX
FH Key Location/Qualifiers
FT CDS 1..2898
FT /*tag= a
FT /product= "Chlamydia trachomatis L2 PMPE protein #1"
XX
PN WO200228898-A2.
XX
PD 11-APR-2002.
XX
PF 28-SEP-2001; 2001WO-US30345.
XX
PR 02-OCT-2000; 2000US-0677752.
XX

PA (ANTE-) ANTEX BIOLOGICS INC.
XX
PI Jackson WJ;
XX
DR WPI; 2002-426107/45.
DR P-PSDB; AAE23453.
XX
PT Novel purified Chlamydia polymorphic membrane protein E or I, useful
PT for preparing vaccines for preventing or treating diseases associated
PT with Chlamydia infection such as trachoma, and infertility -
XX
PS Disclosure; Fig 5; 160pp; English.
XX
CC The invention relates to Chlamydia sp. polymorphic membrane protein
CC (PMP) E and PMPI and nucleic acid molecules encoding such proteins.
CC PMPI and PMPE polypeptides, nucleic acids and vaccines are useful for
CC preventing, treating or ameliorating trachoma, conjunctivitis, tubal
CC occlusion, cervicitis, cervical cancer, pelvic inflammatory disease
CC (PID), lymphogranuloma venereum (LGV), endometritis, epididymitis,
CC salpingitis, infertility, reactive arthritis, inflammatory heart
CC disease, urethritis, autoimmune myocarditis, cardiomyopathy and
CC atherosclerosis. They are also useful as reagents for clinical or
CC medical diagnosis of Chlamydia infections. The present sequence is
CC Chlamydia trachomatis L2 PMPE DNA.
XX
SQ Sequence 2898 BP; 863 A; 625 C; 563 G; 847 T; 0 other;

Query Match 100.0%; Score 2898; DB 24; Length 2898;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2898; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATTGGAAACTCCCTATCAGGACTAGCTAGAGAG 60
Db       1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATTGGAAACTCCCTATCAGGACTAGCTAGAGAG 60

Qy      61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120
Db      61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120

Qy     121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180
Db     121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180

Qy     181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240
Db     181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240

Qy     241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300
Db     241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300

Qy     301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360
Db     301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360

Qy     361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACCATTTACATCG 420
Db     361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACCATTTACATCG 420

Qy     421 AGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 480
Db     421 AGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 480

Qy     481 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCGA 540
Db     481 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCGA 540

Qy     541 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 600
Db     541 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 600

Qy     601 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 660
Db     601 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 660

Qy     661 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTTATCAATAACGCCTGT 720
Db     661 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTTATCAATAACGCCTGT 720

Qy     721 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 780
Db     721 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 780

Qy     781 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 840
Db     781 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 840

Qy     841 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 900
Db     841 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 900

Qy     901 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 960
Db     901 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 960

Qy     961 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1020
Db     961 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1020

Qy    1021 TATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1080
Db    1021 TATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1080

Qy    1081 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1140
Db    1081 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1140

Qy    1141 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1200
Db    1141 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1200

Qy    1201 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 1260
Db    1201 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 1260

Qy    1261 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 1320
Db    1261 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 1320

Qy    1321 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 1380
Db    1321 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 1380

Qy    1381 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 1440
Db    1381 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 1440

Qy    1441 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTGCAGGAAAT 1500
Db    1441 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTGCAGGAAAT 1500

Qy    1501 TCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTG 1560
Db    1501 TCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTG 1560

Qy    1561 AAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTAT 1620
Db    1561 AAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTAT 1620

Qy    1621 ACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGAC 1680
Db    1621 ACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGAC 1680

Qy    1681 TATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTGTCATCACAGCCTATG 1740
Db    1681 TATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTGTCATCACAGCCTATG 1740
```

```
Qy   1741 CTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGATATGGATTTTTCGGGA 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 CTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGATATGGATTTTTCGGGA 1800

Qy   1801 CTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 CTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAA 1860

Qy   1861 GATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAAGCCAATAGATTCCAT 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAAGCCAATAGATTCCAT 1920

Qy   1921 AGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGCCCGAAACACAGAAGT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 AGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGCCCGAAACACAGAAGT 1980

Qy   1981 CCCCTCATAGCGAATACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAAT 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 CCCCTCATAGCGAATACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAAT 2040

Qy   2041 AGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 AGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 2100

Qy   2101 ATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGA 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 ATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGA 2160

Qy   2161 TACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTGAAATTCAGTCAGACC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 TACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTGAAATTCAGTCAGACC 2220

Qy   2221 TACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 TACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGC 2280

Qy   2281 CAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTT 2340

Qy   2341 TACAGCTATGGAGACCATAACTGTCACCATTTCTATACCCAAGGAGAAAATCTAACATCT 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 TACAGCTATGGAGACCATAACTGTCACCATTTCTATACCCAAGGAGAAAATCTAACATCT 2400

Qy   2401 CAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTGTTTTTTTTGATCTCCCTATGAAA 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 CAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTGTTTTTTTTGATCTCCCTATGAAA 2460

Qy   2461 CCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 CCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCT 2520

Qy   2521 AGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCT 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 AGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCT 2580

Qy   2581 TTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCAAAGA 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 TTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCAAAGA 2640

Qy   2641 CCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACCAGGG 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 CCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACCAGGG 2700

Qy   2701 ATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCG 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 ATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCG 2760

Qy   2761 CGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTC 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 CGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTC 2820

Qy   2821 CATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAA 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 CATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAA 2880

Qy   2881 ATTGCTCTGCGATTCTAG 2898
          ||||||||||||||||||
Db   2881 ATTGCTCTGCGATTCTAG 2898
```

RESULT 2
AAZ01425
ID AAZ01425 standard; DNA; 1038602 BP.
XX
AC AAZ01425;
XX
DT 07-OCT-1999 (first entry)
XX
DE Complete genome sequence of Chlamydia trachomatis.
XX
KW Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;
KW nongonococcal uretritis; epidymitis; cervicitis; salpingitis;
KW bartholinitis; pneumopathy; venereal lymphogranulomatosis; ss.
XX
OS Chlamydia trachomatis.
XX
PN WO9928475-A2.
XX
PD 10-JUN-1999.
XX
PF 27-NOV-1998; 98WO-IB01939.
XX
PR 04-NOV-1998; 98US-0107077.
PR 28-NOV-1997; 97FR-0015041.
PR 17-DEC-1997; 97FR-0016034.
XX
PA (GEST) GENSET.
XX
PI Griffais R;
XX
DR WPI; 1999-371125/31.
XX
PT Genome sequence of Chlamydia trachomatis
XX
PS Claim 1; Page 373-656; 1755pp; English.
XX
CC The present sequence represents the complete genome of Chlamydia
CC trachomatis. Open reading frames (ORFs) of the genome encode
CC polypeptides AAY36754-Y37949. The polypeptides can be used as vaccines
CC against Chlamydia trachomatis. Antisense and ribozyme sequences can also
CC be used to control growth of the microorganism. Chlamydia trachomatis is
CC responsible for a large number of diseases, e.g. eye diseases such as
CC conventional trachoma, nonendemic trachoma, paratrachoma, and inclusion
CC conjunctivitis; genital diseases such as nongonococcal uretritis,
CC epidymitis, cervicitis, salpingitis, perihepatitis, bartholinitis;
CC pneumopathy in breast feeding infants; and venereal
CC lymphogranulomatosis. The polypeptides of the invention may be of use in
CC treating these diseases.
XX
SQ Sequence 1038602 BP; 304265 A; 214645 C; 214259 G; 305001 T; 432 other;

Query Match 98.9%; Score 2866; DB 20; Length 1038602;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2878; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```
Qy      1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATTGGAAACTCCCTATCAGGACTAGCTAGAGAG 60
          |||||||||||| || || |||||||||||||||| ||||||||||||||||||||||||
Db 484964 ATGAAAAAAGCATTCTTTTTTTTCCTTATTGGAAGCTCCCTATCAGGACTAGCTAGAGAG 485023

Qy     61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 485024 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 485083
```

```
Qy     121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 485084 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 485143

Qy     181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 485144 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 485203

Qy     241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 485204 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 485263

Qy     301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 485264 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 485323

Qy     361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACCATTTACATCG 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 485324 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACCATTTACATCG 485383

Qy     421 AGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 485384 AGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCAATTCATGCTCAAAAT 485443

Qy     481 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCGA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 485444 CTTTACATAAATCACAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCGA 485503

Qy     541 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 485504 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 485563

Qy     601 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 485564 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 485623

Qy     661 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTTATCAATAACGCCTGT 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 485624 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTTATCAATAACGCCTGT 485683

Qy     721 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 485684 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 485743

Qy     781 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 485744 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAATCTTCTGAT 485803

Qy     841 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 485804 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 485863

Qy     901 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 485864 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 485923

Qy     961 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 485924 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 485983

Qy    1021 TATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 485984 TATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 486043

Qy    1081 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 486044 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 486103

Qy    1141 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 486104 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 486163

Qy    1201 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 486164 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 486223

Qy    1261 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 486224 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 486283

Qy    1321 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 486284 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 486343

Qy    1381 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 486344 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 486403

Qy    1441 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTGCAGGAAAT 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 486404 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTGCAGGAAAT 486463

Qy    1501 TCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTG 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 486464 TCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTG 486523

Qy    1561 AAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTAT 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 486524 AAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTAT 486583

Qy    1621 ACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGAC 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 486584 ACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGAC 486643

Qy    1681 TATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTGTCATCACAGCCTATG 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 486644 TATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTGTCATCACAGCCTATG 486703

Qy    1741 CTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGATATGGATTTTTCGGGA 1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 486704 CTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGATATGGATTTCTCGGGA 486763

Qy    1801 CTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAA 1860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 486764 CTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGAGTTGGGGCTGGGCAAAAACTCAA 486823

Qy    1861 GATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAAGCCAATAGATTCCAT 1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 486824 GATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCCAAAAAAGCCAATAGATTCCAT 486883

Qy    1921 AGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGCCCGAAACACAGAAGT 1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 486884 AGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGCCCGAAACACAGAAGT 486943

Qy    1981 CCCCTCATAGCGAATACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAAT 2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 486944 CCCCTCATAGCGAATACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAAT 487003

Qy    2041 AGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 487004 AGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 487063

Qy    2101 ATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGA 2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 487064 ATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGA 487123

Qy    2161 TACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTGAAATTCAGTCAGACC 2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 487124 TACTCTGCGGGGATGATAGCAGGGCAAACACATACCTTCTCATTGAAATTCAGTCAGACC 487183

Qy    2221 TACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGC 2280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 487184 TACACCAAACTCAATGAGCGTTACGCGAAAAACAACGTATCTTCTAAAAATTACTCATGC 487243

Qy    2281 CAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTT 2340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 487244 CAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGGCTAAATTAGTTGGTCTT 487303

Qy   2341 TACAGCTATGGAGACCATAACTGTCACCATTTCTATACCCAAGGAGAAAATCTAACATCT 2400
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db 487304 TACAGCTATGGAGATCATAACTGTCACCATTTCTATACCCAAGGAGAAAATCTAACATCT 487363

Qy   2401 CAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTGTTTTTTTTTGATCTCCCTATGAAA 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 487364 CAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTGTTTTTTTTTGATCTCCCTATGAAA 487423

Qy   2461 CCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 487424 CCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCT 487483

Qy   2521 AGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCT 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 487484 AGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCT 487543

Qy   2581 TTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCAAAGA 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 487544 TTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCAAAGA 487603

Qy   2641 CCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACCAGGG 2700
          |||||||||||||||||||||||||||||||||||||||||||||||||||| || | |
Db 487604 CCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACTAGAG 487663

Qy   2701 ATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCG 2760
          ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db 487664 ATCGCGACCCAGCTCCTAGCCAGTAAAGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCG 487723

Qy   2761 CGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTC 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 487724 CGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTC 487783

Qy   2821 CATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAA 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 487784 CATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAA 487843

Qy   2881 ATTGCTCTGCGATTCTAG 2898
          ||||||||||||||||||
Db 487844 ATTGCTCTGCGATTCTAG 487861


RESULT 3
AAH56256
ID   AAH56256 standard; DNA; 2895 BP.
XX
AC   AAH56256;
XX
DT   05-SEP-2001  (first entry)
XX
DE   Chlamydia trachomatis pmpE gene.
XX
KW   Chlamydia; vaccine; infection; fusion protein; antigen;
KW   pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;
KW   acute respiratory tract infection; Cap1; CT529; OMCB;
KW   polymorphic membrane protein; pmp; thiol specific antioxidant; TSA; ds.
XX
OS   Chlamydia trachomatis.
XX
PN   WO200140474-A2.
XX
PD   07-JUN-2001.
XX
PF   04-DEC-2000; 2000WO-US32919.
XX
PR   03-DEC-1999;   99US-0454684.
PR   19-APR-2000; 2000US-0556877.
PR   20-JUN-2000; 2000US-0598419.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Probst P,  Bhatia A,  Skeiky YAW,  Fling SP,  Scholler J;
XX
DR   WPI; 2001-374831/39.
XX
PT   Chlamydia polypeptides and fusion proteins useful for preventing pelvic
PT   inflammatory disease, trachoma, acute respiratory tract infections,
PT   atherosclerosis and heart disease -
XX
PS   Claim 1; Page 181-182; 295pp; English.
XX
CC   The present nucleotide sequence is provided in a specification
CC   relating to compounds and methods for the treatment and diagnosis of
CC   chlamydial infection. The compounds provided include polypeptides and
CC   fusion proteins comprising immunogenic portions of Chlamydia antigens
CC   and DNA sequences encoding such polypeptides. They are useful for
CC   vaccinating against chlamydial infection, which causes pelvic
CC   inflammatory disease, trachoma, acute respiratory tract infections,
CC   atherosclerosis and heart disease.
XX
SQ   Sequence 2895 BP; 866 A; 632 C; 558 G; 839 T; 0 other;

  Query Match             97.3%;  Score 2821.2;  DB 22;  Length 2895;
  Best Local Similarity   98.6%;  Pred. No. 0;
  Matches 2857;  Conservative    0;  Mismatches   38;  Indels    3;  Gaps    1;

Qy      1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATTGGAAACTCCCTATCAGGACTAGCTAGAGAG 60
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db      1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATCGGAAACTCCCTATCAGGACTAGCTAGAGAG 60

Qy     61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120

Qy    121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180

Qy    181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240

Qy    241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300

Qy    301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360

Qy    361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACCATTTACATCG 420
          |||||||||||||||||||||||||||||||||||||||||||||| |||||||| |
Db    361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACTATTTACATGG 420

Qy    421 AGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 480
          || |||||| |||| ||||   ||||||||||||||||||||||||||||||||||||||
Db    421 AGAAATCCTTATGCTGCTG---ATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 477

Qy    481 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCGA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    478 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCAA 537

Qy    541 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    538 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 597

Qy    601 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    598 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 657

Qy    661 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTTATCAATAACGCCTGT 720
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    658 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTCATCAATAACGCCTGT 717

Qy    721 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 780
```

```
Db    718 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 777

Qy    781 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    778 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 837

Qy    841 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    838 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 897

Qy    901 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    898 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 957

Qy    961 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    958 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1017

Qy   1021 TATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1080
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   1018 TATATAGACGGAACCAGTAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1077

Qy   1081 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1078 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1137

Qy   1141 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1138 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1197

Qy   1201 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1198 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 1257

Qy   1261 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1258 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 1317

Qy   1321 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1318 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 1377

Qy   1381 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1378 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 1437

Qy   1441 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTGCAGGAAAT 1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||| ||
Db   1438 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTACAGGAGAT 1497

Qy   1501 TCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTG 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1498 TCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTG 1557

Qy   1561 AAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTAT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1558 AAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTAT 1617

Qy   1621 ACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGAC 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1618 ACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGAC 1677

Qy   1681 TATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTGTCATCACAGCCTATG 1740
          || ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1678 TACGGGAACTCTCCTTATGAATCCACAGATCTGACCCATGCTCTGTCATCACAGCCTATG 1737

Qy   1741 CTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGATATGGATTTTTCGGGA 1800
          ||||||||||||||| ||||| |||||||||||  ||| ||  |||| ||||||||||||
Db   1738 CTATCTATTTCTGAAGCTAGCGATAACCAGCTACAATCAGAAAATATAGATTTTTCGGGA 1797

Qy   1801 CTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1798 CTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAA 1857

Qy   1861 GATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAAGCCAATAGATTCCAT 1920
          |||||||||||||||||||||||||||||| |||||||||||||||||||||| |||
Db   1858 GATCCAGAACCAGCATCTTCAGCAACAATCACTGATCCACAAAAAGCCAATAGATTTCAT 1917

Qy   1921 AGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGCCCGAAACACAGAAGT 1980
          ||||||||| |||| || |||||||||||| |||||||||||||||||| ||||||||||
Db   1918 AGAACCTTACTACTAACATGGCTTCCTGCCGGGTATGTTCCTAGCCCAAAACACAGAAGT 1977

Qy   1981 CCCCTCATAGCGAATACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAAT 2040
          |||||||||| || ||||||||||||||||||||||||||||||||||||||||||||||
Db   1978 CCCCTCATAGCTAACACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAAT 2037

Qy   2041 AGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 2100
          |||||||| |||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   2038 AGTGCAGAGCTGACACCTAGTGGTCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 2097

Qy   2101 ATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGA 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2098 ATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGA 2157

Qy   2161 TACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTGAAATTCAGTCAGACC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2158 TACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTGAAATTCAGTCAGACC 2217

Qy   2221 TACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2218 TACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGC 2277

Qy   2281 CAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2278 CAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTT 2337

Qy   2341 TACAGCTATGGAGACCATAACTGTCACCATTTCTATACCCAAGGAGAAAATCTAACATCT 2400
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   2338 TACAGCTATGGAGACCATAACTGTCACCATTTCTATACTCAAGGAGAAAATCTAACATCT 2397

Qy   2401 CAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTGTTTTTTTTGATCTCCCTATGAAA 2460
          ||||||||||||||| |||||||||||||||||||| |||||||||||||||||||||||
Db   2398 CAAGGGACGTTCCGCAGTCAAACGATGGGAGGTGCTGTCTTTTTTGATCTCCCTATGAAA 2457

Qy   2461 CCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2458 CCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCT 2517

Qy   2521 AGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCT 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2518 AGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCT 2577

Qy   2581 TTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCAAAGA 2640
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   2578 TTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCACAGA 2637

Qy   2641 CCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACCAGGG 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2638 CCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACCAGGG 2697

Qy   2701 ATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCG 2760
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   2698 ATCGCGACCCAGCTCCTAGCCAGTAAAGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCG 2757

Qy   2761 CGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTC 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2758 CGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTC 2817

Qy   2821 CATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAA 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2818 CATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAA 2877

Qy   2881 ATTGCTCTGCGATTCTAG 2898
          ||||||||||||||||||
Db   2878 ATTGCTCTGCGATTCTAG 2895
```

RESULT 4
ABL92485
ID ABL92485 standard; DNA; 2895 BP.
XX
AC ABL92485;
XX
DT 05-JUN-2002 (first entry)
XX
DE Chlamydia DNA sequence SEQ ID NO:171.
XX
KW Chlamydial infection; Chlamydia; vaccine; detection; diagnosis;
KW antigen; antibacterial; immunostimulant; immune response;
KW Chlamydia-specific T-cell response; gene; ds.
XX
OS Chlamydia sp.
XX
PN WO200208267-A2.
XX
PD 31-JAN-2002.
XX
PF 20-JUL-2001; 2001WO-US23121.
XX
PR 20-JUL-2000; 2000US-0620412.
PR 23-APR-2001; 2001US-0841132.
XX
PA (CORI-) CORIXA CORP.
XX
PI Fling SP, Skeiky YAW, Probst P, Bhatia A;
XX
DR WPI; 2002-179901/23.
XX
PT Novel compositions comprising Chlamydia Capl protein and its use in the
PT treatment of Chlamydia infection -
XX
PS Example 1; Page 214; 537pp; English.
XX
CC The present invention describes compositions comprising a Chlamydia Capl
CC protein and methods for the diagnosis and therapy of Chlamydia infection.
CC Chlamydia DNA and protein sequences from the present invention can have
CC antibacterial and immunostimulant activities, and can be used in
CC vaccines. Compounds from the present invention can be used for eliciting
CC an immune response, specifically stimulating a Chlamydia-specific T-cell
CC response or inhibiting the development of a Chlamydia infection in an
CC animal. Methods from the present invention can be used: for detecting the
CC presence of Chlamydia in a patient; to stimulate and/or expand T cells
CC specific for a Chlamydia protein; and for treatment of a Chlamydia
CC infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 2895 BP; 866 A; 632 C; 558 G; 839 T; 0 other;

Query Match 97.3%; Score 2821.2; DB 24; Length 2895;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 2857; Conservative 0; Mismatches 38; Indels 3; Gaps 1;

```
Qy      1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATTGGAAACTCCCTATCAGGACTAGCTAGAGAG 60
          ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db      1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATCGGAAACTCCCTATCAGGACTAGCTAGAGAG 60

Qy     61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120

Qy    121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180

Qy    181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240

Qy    241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300

Qy    301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360

Qy    361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACCATTTACATCG 420
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||  |
Db    361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACTATTTACATGG 420

Qy    421 AGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 480
          ||  ||||||  ||||  ||||   ||||||||||||||||||||||||||||||||||||
Db    421 AGAAATCCTTATGCTGCTG---ATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 477

Qy    481 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCGA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    478 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCAA 537

Qy    541 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    538 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 597

Qy    601 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    598 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 657

Qy    661 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTTATCAATAACGCCTGT 720
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    658 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTCATCAATAACGCCTGT 717

Qy    721 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    718 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 777

Qy    781 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    778 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 837

Qy    841 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    838 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 897

Qy    901 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    898 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 957

Qy    961 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    958 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1017

Qy   1021 TATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1080
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   1018 TATATAGACGGAACCAGTAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1077

Qy   1081 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1078 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1137

Qy   1141 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1138 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1197

Qy   1201 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1198 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 1257

Qy   1261 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1258 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 1317

Qy   1321 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 1380
```

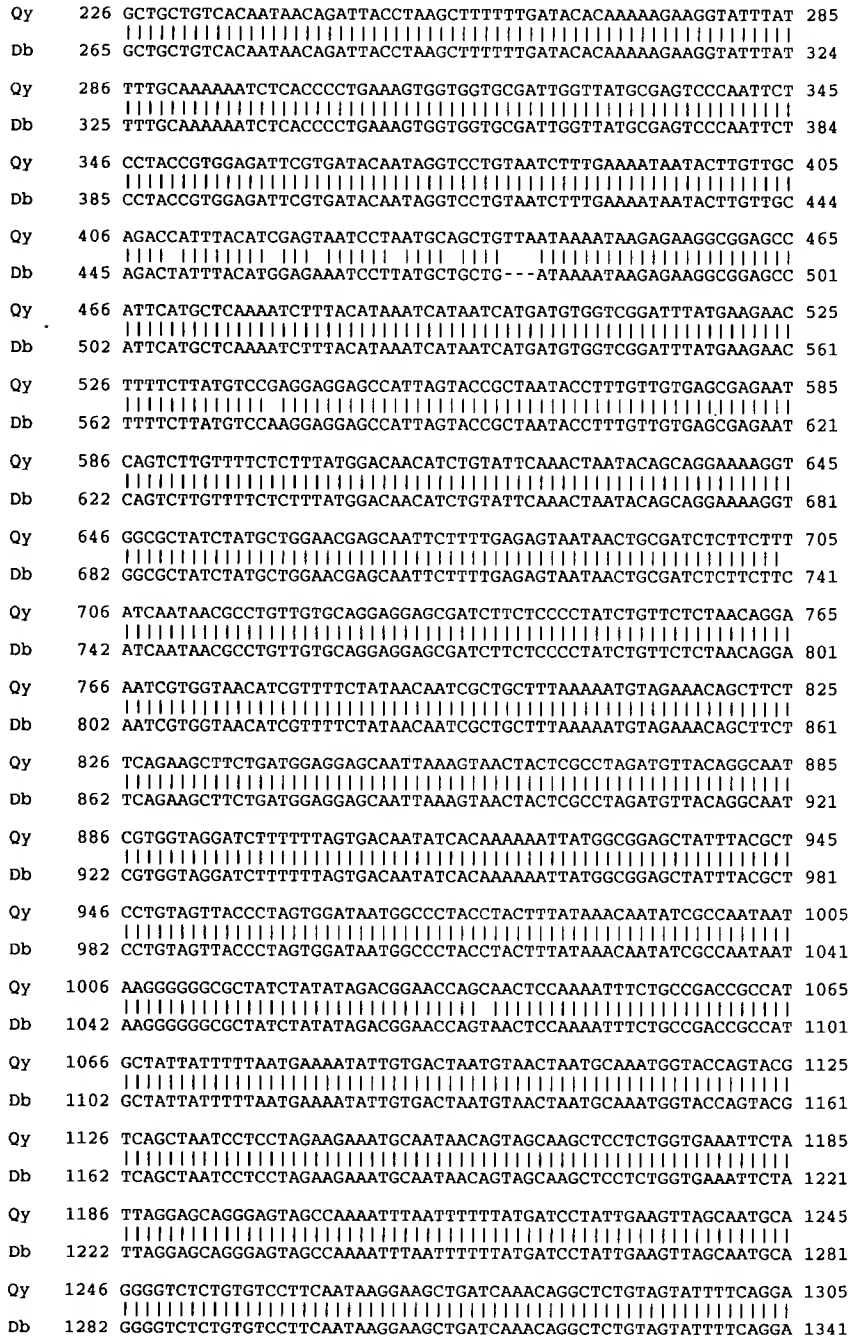
```
Db   1318 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 1377

Qy   1381 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 1440

Db   1378 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 1437

Qy   1441 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTGCAGGAAAT 1500

Db   1438 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTACAGGAGAT 1497

Qy   1501 TCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTG 1560

Db   1498 TCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTG 1557

Qy   1561 AAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTAT 1620

Db   1558 AAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTAT 1617

Qy   1621 ACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGAC 1680

Db   1618 ACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGAC 1677

Qy   1681 TATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTGTCATCACAGCCTATG 1740

Db   1678 TACGGGAACTCTCCTTATGAATCCACAGATCTGACCCATGCTCTGTCATCACAGCCTATG 1737

Qy   1741 CTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGATATGGATTTTTCGGGA 1800

Db   1738 CTATCTATTTCTGAAGCTAGCGATAACCAGCTACAATCAGAAAATATAGATTTTTCGGGA 1797

Qy   1801 CTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAA 1860

Db   1798 CTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAA 1857

Qy   1861 GATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAAGCCAATAGATTCCAT 1920

Db   1858 GATCCAGAACCAGCATCTTCAGCAACAATCACTGATCCACAAAAAGCCAATAGATTTCAT 1917

Qy   1921 AGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGCCCGAAACACAGAAGT 1980

Db   1918 AGAACCTTACTACTAACATGGCTTCCTGCCGGGTATGTTCCTAGCCCAAAACACAGAAGT 1977

Qy   1981 CCCCTCATAGCGAATACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAAT 2040

Db   1978 CCCCTCATAGCTAACACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAAT 2037

Qy   2041 AGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 2100

Db   2038 AGTGCAGAGCTGACACCTAGTGGTCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 2097

Qy   2101 ATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGA 2160

Db   2098 ATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGA 2157

Qy   2161 TACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTGAAATTCAGTCAGACC 2220

Db   2158 TACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTGAAATTCAGTCAGACC 2217

Qy   2221 TACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGC 2280

Db   2218 TACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGC 2277

Qy   2281 CAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTT 2340

Db   2278 CAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTT 2337

Qy   2341 TACAGCTATGGAGACCATAACTGTCACCATTTCTATACCCAAGGAGAAAATCTAACATCT 2400

Db   2338 TACAGCTATGGAGACCATAACTGTCACCATTTCTATACTCAAGGAGAAAATCTAACATCT 2397

Qy   2401 CAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTGTTTTTTTTGATCTCCCTATGAAA 2460

Db   2398 CAAGGGACGTTCCGCAGTCAAACGATGGGAGGTGCTGTCTTTTTTGATCTCCCTATGAAA 2457

Qy   2461 CCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCT 2520

Db   2458 CCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCT 2517

Qy   2521 AGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCT 2580

Db   2518 AGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCT 2577

Qy   2581 TTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCAAAGA 2640

Db   2578 TTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCACAGA 2637

Qy   2641 CCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACCAGGG 2700

Db   2638 CCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACCAGGG 2697

Qy   2701 ATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCG 2760

Db   2698 ATCGCGACCCAGCTCCTAGCCAGTAAAGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCG 2757

Qy   2761 CGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTC 2820

Db   2758 CGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTC 2817

Qy   2821 CATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAA 2880

Db   2818 CATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAA 2877

Qy   2881 ATTGCTCTGCGATTCTAG 2898

Db   2878 ATTGCTCTGCGATTCTAG 2895
```

RESULT 5
AAA64753
ID AAA64753 standard; DNA; 2895 BP.
XX
AC AAA64753;
XX
DT 02-FEB-2001 (first entry)
XX
DE C. trachomatis pmpE gene coding sequence.
XX
KW Chlamydial infection; sexually transmitted disease;
KW pelvic inflammatory disease; PID; tubal obstruction; infertility;
KW trachoma; blindness; acute respiratory tract infection;
KW atherosclerosis; coronary heart disease; antibacterial; ss.
XX
OS Chlamydia trachomatis.
XX
PN WO200034483-A2.
XX
PD 15-JUN-2000.
XX
PF 08-DEC-1999; 99WO-US29012.
XX
PR 08-DEC-1998; 98US-0208277.
PR 08-APR-1999; 99US-0288594.
PR 01-OCT-1999; 99US-0410568.
PR 22-OCT-1999; 99US-0426571.
XX
PA (CORI-) CORIXA CORP.
XX
PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
XX
DR WPI; 2000-431303/37.
XX
PT Isolated polypeptide for diagnosis and treatment of Chlamydia infection
PT comprises immunogenic portion of Chlamydia antigen, which comprises
PT amino acid sequence encoded by polynucleotide sequence -
XX

```
PS   Claim 1; Pages 173-174; 256pp; English.
XX
CC   The present invention relates to new nucleic acid sequences and the
CC   proteins encoded by the nucleic acid sequences. The encoded proteins
CC   comprise an immunogenic portion of a Chlamydia antigen. The encoded
CC   proteins are useful for the serodiagnosis and treatment of Chlamydia
CC   infection. Chlamdiae are intracellular bacterial pathogens that are
CC   responsible for a wide variety of human infections. C. trachomatis
CC   infection is one of the most common sexually transmitted diseases and can
CC   lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
CC   and infertility. Trachoma due to ocular infection with C. trachomatis is
CC   the leading cause of preventable blindness worldwide. C. pneumonia is a
CC   major cause of acute respiratory tract infections in humans and is also
CC   thought to play a role in the pathogenesis of atherosclerosis and
CC   coronary heart disease. The present sequence is a nucleic acid sequence
CC   isolated in the present invention.
XX
SQ   Sequence 2895 BP; 866 A; 633 C; 558 G; 838 T; 0 other;

  Query Match             97.3%;  Score 2819.6;  DB 21;  Length 2895;
  Best Local Similarity   98.6%;  Pred. No. 0;
  Matches 2856;  Conservative    0;  Mismatches   39;  Indels    3;  Gaps    1;

Qy      1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATTGGAAACTCCCTATCAGGACTAGCTAGAGAG 60
          ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db      1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATCGGAAACTCCCTATCAGGACTAGCTAGAGAG 60

Qy     61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120

Qy    121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180

Qy    181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240

Qy    241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCCC 300

Qy    301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360

Qy    361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACCATTTACATCG 420
          |||||||||||||||||||||||||||||||||||||||||||||||| |||||||| |
Db    361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACTATTTACATGG 420

Qy    421 AGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 480
          || |||||| |||| ||||   ||||||||||||||||||||||||||||||||||||||
Db    421 AGAAATCCTTATGCTGCTG---ATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 477

Qy    481 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCGA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    478 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCAA 537

Qy    541 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    538 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 597

Qy    601 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    598 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 657

Qy    661 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTTATCAATAACGCCTGT 720
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    658 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTCATCAATAACGCCTGT 717

Qy    721 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    718 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 777

Qy    781 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    778 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 837

Qy    841 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    838 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 897

Qy    901 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    898 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 957

Qy    961 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    958 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1017

Qy   1021 TATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1080
          ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   1018 TATATAGACGGAACCAGTAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1077

Qy   1081 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1078 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1137

Qy   1141 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1138 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1197

Qy   1201 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1198 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 1257

Qy   1261 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1258 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 1317

Qy   1321 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1318 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 1377

Qy   1381 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1378 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 1437

Qy   1441 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTGCAGGAAAT 1500
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||| ||
Db   1438 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTACAGGAGAT 1497

Qy   1501 TCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTG 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1498 TCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTG 1557

Qy   1561 AAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTAT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1558 AAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTAT 1617

Qy   1621 ACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGAC 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1618 ACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGAC 1677

Qy   1681 TATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTGTCATCACAGCCTATG 1740
          || ||||| |||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   1678 TACGGGAACTCTCCTTATGAATCCACAGATCTGACCCATGCTCTGTCATCACAGCCTATG 1737

Qy   1741 CTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGATATGGATTTTTCGGGA 1800
          |||||||||||||| |||||||||||||  ||| ||  |||| ||||||||||||||
Db   1738 CTATCTATTTCTGAAGCTAGCGATAACCAGCTACAATCAGAAAATATAGATTTTTCGGGA 1797

Qy   1801 CTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1798 CTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAA 1857
```

```
Qy   1861 GATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAAGCCAATAGATTCCAT 1920
          ||||||||||||||||||||||||||||||| |||||||||||||||||||||||| |||
Db   1858 GATCCAGAACCAGCATCTTCAGCAACAATCACTGATCCACAAAAAGCCAATAGATTTCAT 1917

Qy   1921 AGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGCCCGAAACACAGAAGT 1980
          ||||||||| |||| || |||||||||| |||||||||||||||| ||||||||||||
Db   1918 AGAACCTTACTACTAACATGGCTTCCTGCCGGGTATGTTCCTAGCCCAAAACACAGAAGT 1977

Qy   1981 CCCCTCATAGCGAATACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAAT 2040
          |||||||||| || ||||||||||||||||||||||||||||||||||||||||||||||
Db   1978 CCCCTCATAGCTAACACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAAT 2037

Qy   2041 AGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 2100
          |||||||| |||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   2038 AGTGCAGAGCTGACACCTAGTGGTCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 2097

Qy   2101 ATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGA 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2098 ATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGA 2157

Qy   2161 TACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTGAAATTCAGTCAGACC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2158 TACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTGAAATTCAGTCAGACC 2217

Qy   2221 TACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2218 TACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGC 2277

Qy   2281 CAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2278 CAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTT 2337

Qy   2341 TACAGCTATGGAGACCATAACTGTCACCATTTCTATACCCAAGGAGAAAATCTAACATCT 2400
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   2338 TACAGCTATGGAGACCATAACTGTCACCATTTCTATACTCAAGGAGAAAATCTAACATCT 2397

Qy   2401 CAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTGTTTTTTTTGATCTCCCTATGAAA 2460
          |||||||||||||| |||||||||||||||||||| ||||||||||||||||||||||||
Db   2398 CAAGGGACGTTCCGCAGTCAAACGATGGGAGGTGCTGTCTTTTTTGATCTCCCTATGAAA 2457

Qy   2461 CCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2458 CCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCT 2517

Qy   2521 AGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCT 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2518 AGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCT 2577

Qy   2581 TTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCAAAGA 2640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   2578 TTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCACAGA 2637

Qy   2641 CCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACCAGGG 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2638 CCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACCAGGG 2697

Qy   2701 ATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCG 2760
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   2698 ATCGCGACCCAGCTCCTAGCCAGTAAAGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCG 2757

Qy   2761 CGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTC 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2758 CGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTC 2817

Qy   2821 CATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAA 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2818 CATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAA 2877

Qy   2881 ATTGCTCTGCGATTCTAG 2898
          ||||||||||||||||||
Db   2878 ATTGCTCTGCGATTCTAG 2895
```

```
RESULT 6
AAA64759
ID   AAA64759 standard; DNA; 2934 BP.
XX
AC   AAA64759;
XX
DT   02-FEB-2001  (first entry)
XX
DE   C. trachomatis pmpE gene coding sequence minus the signal sequence.
XX
KW   Chlamydial infection; sexually transmitted disease;
KW   pelvic inflammatory disease; PID; tubal obstruction; infertility;
KW   trachoma; blindness; acute respiratory tract infection;
KW   atherosclerosis; coronary heart disease; antibacterial; ss.
XX
OS   Chlamydia trachomatis.
XX
PN   WO200034483-A2.
XX
PD   15-JUN-2000.
XX
PF   08-DEC-1999;   99WO-US29012.
XX
PR   08-DEC-1998;   98US-0208277.
PR   08-APR-1999;   99US-0288594.
PR   01-OCT-1999;   99US-0410568.
PR   22-OCT-1999;   99US-0426571.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Probst P,  Bhatia A,  Skeiky YAW,  Fling SP,  Jen S,  Stromberg EJ;
XX
DR   WPI; 2000-431303/37.
XX
PT   Isolated polypeptide for diagnosis and treatment of Chlamydia infection
PT   comprises immunogenic portion of Chlamydia antigen, which comprises
PT   amino acid sequence encoded by polynucleotide sequence  -
XX
PS   Claim 1; Pages 200-201; 256pp; English.
XX
CC   The present invention relates to new nucleic acid sequences and the
CC   proteins encoded by the nucleic acid sequences. The encoded proteins
CC   comprise an immunogenic portion of a Chlamydia antigen. The encoded
CC   proteins are useful for the serodiagnosis and treatment of Chlamydia
CC   infection. Chlamdiae are intracellular bacterial pathogens that are
CC   responsible for a wide variety of human infections. C. trachomatis
CC   infection is one of the most common sexually transmitted diseases and can
CC   lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
CC   and infertility. Trachoma due to ocular infection with C. trachomatis is
CC   the leading cause of preventable blindness worldwide. C. pneumonia is a
CC   major cause of acute respiratory tract infections in humans and is also
CC   thought to play a role in the pathogenesis of atherosclerosis and
CC   coronary heart disease. The present sequence is a nucleic acid sequence
CC   isolated in the present invention.
XX
SQ   Sequence 2934 BP; 875 A; 645 C; 575 G; 839 T; 0 other;

  Query Match             95.9%;  Score 2777.8;  DB 21;  Length 2934;
  Best Local Similarity   98.6%;  Pred. No. 0;
  Matches 2813;  Conservative    0;  Mismatches   37;  Indels    3;  Gaps    1;

Qy      46 GGACTAGCTAGAGAGGTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCT 105
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85 GGACTAGCTAGAGAGGTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCT 144

Qy     106 ACGAAAGAGTCGCTATCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAAC 165
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     145 ACGAAAGAGTCGCTATCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAAC 204

Qy     166 TGCTATCTCGATAACCTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGA 225
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     205 TGCTATCTCGATAACCTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGA 264
```


Qy 226 GCTGCTGTCACAATAACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTAT 285
Db 265 GCTGCTGTCACAATAACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTAT 324
Qy 286 TTTGCAAAAAATCTCACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCT 345
Db 325 TTTGCAAAAAATCTCACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCT 384
Qy 346 CCTACCGTGGAGATTCGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGC 405
Db 385 CCTACCGTGGAGATTCGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGC 444
Qy 406 AGACCATTTACATCGAGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCC 465
Db 445 AGACTATTTACATGGAGAAATCCTTATGCTGCTG---ATAAAATAAGAGAAGGCGGAGCC 501
Qy 466 ATTCATGCTCAAAATCTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAAC 525
Db 502 ATTCATGCTCAAAATCTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAAC 561
Qy 526 TTTTCTTATGTCCGAGGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAAT 585
Db 562 TTTTCTTATGTCCAAGGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAAT 621
Qy 586 CAGTCTTGTTTTCTCTTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGT 645
Db 622 CAGTCTTGTTTTCTCTTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGT 681
Qy 646 GGCGCTATCTATGCTGGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTT 705
Db 682 GGCGCTATCTATGCTGGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTC 741
Qy 706 ATCAATAACGCCTGTTGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGA 765
Db 742 ATCAATAACGCCTGTTGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGA 801
Qy 766 AATCGTGGTAACATCGTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCT 825
Db 802 AATCGTGGTAACATCGTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCT 861
Qy 826 TCAGAAGCTTCTGATGGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAAT 885
Db 862 TCAGAAGCTTCTGATGGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAAT 921
Qy 886 CGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCT 945
Db 922 CGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCT 981
Qy 946 CCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAAT 1005
Db 982 CCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAAT 1041
Qy 1006 AAGGGGGGCGCTATCTATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCAT 1065
Db 1042 AAGGGGGGCGCTATCTATATAGACGGAACCAGTAACTCCAAAATTTCTGCCGACCGCCAT 1101
Qy 1066 GCTATTATTTTTAATGAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACG 1125
Db 1102 GCTATTATTTTTAATGAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACG 1161
Qy 1126 TCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTA 1185
Db 1162 TCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTA 1221
Qy 1186 TTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCA 1245
Db 1222 TTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCA 1281
Qy 1246 GGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGA 1305
Db 1282 GGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGA 1341
Qy 1306 GCTACTGTTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCC 1365
Db 1342 GCTACTGTTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCC 1401
Qy 1366 CTTACTCTCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGA 1425
Db 1402 CTTACTCTCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGA 1461
Qy 1426 TTCACACAAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAA 1485
Db 1462 TTCACACAAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAA 1521
Qy 1486 AATGGTGCAGGAAATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAAT 1545
Db 1522 AATGGTACAGGAGATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAAT 1581
Qy 1546 CTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAAT 1605
Db 1582 CTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAAT 1641
Qy 1606 AACAGCAATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTC 1665
Db 1642 AACAGCAATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTC 1701
Qy 1666 TCACTCATTGATGACTATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTG 1725
Db 1702 TCACTCATTGATGACTACGGGAACTCTCCTTATGAATCCACAGATCTGACCCATGCTCTG 1761
Qy 1726 TCATCACAGCCTATGCTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGAT 1785
Db 1762 TCATCACAGCCTATGCTATCTATTTCTGAAGCTAGCGATAACCAGCTACAATCAGAAAAT 1821
Qy 1786 ATGGATTTTTCGGGACTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGC 1845
Db 1822 ATAGATTTTTCGGGACTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGC 1881
Qy 1846 TGGGCAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAA 1905
Db 1882 TGGGCAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACTGATCCACAAAAA 1941
Qy 1906 GCCAATAGATTCCATAGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGC 1965
Db 1942 GCCAATAGATTTCATAGAACCTTACTACTAACATGGCTTCCTGCCGGGTATGTTCCTAGC 2001
Qy 1966 CCGAAACACAGAAGTCCCCTCATAGCGAATACCTTATGGGGGAATATGCTGCTTGCAACA 2025
Db 2002 CCAAAACACAGAAGTCCCCTCATAGCTAACACCTTATGGGGGAATATGCTGCTTGCAACA 2061
Qy 2026 GAAAGCTTAAAAAATAGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTACA 2085
Db 2062 GAAAGCTTAAAAAATAGTGCAGAGCTGACACCTAGTGGTCATCCTTTCTGGGGAATTACA 2121
Qy 2086 GGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCAT 2145
Db 2122 GGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCAT 2181
Qy 2146 ATGCGCTCTTCCGGATACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTG 2205
Db 2182 ATGCGCTCTTCCGGATACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTG 2241
Qy 2206 AAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCT 2265
Db 2242 AAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCT 2301
Qy 2266 AAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACT 2325
Db 2302 AAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACT 2361
Qy 2326 AAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCACCATTTCTATACCCAAGGA 2385
Db 2362 AAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCACCATTTCTATACTCAAGGA 2421
Qy 2386 GAAAATCTAACATCTCAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTGTTTTTTTT 2445

```
Db   2422 GAAAATCTAACATCTCAAGGGACGTTCCGCAGTCAAACGATGGGAGGTGCTGTCTTTTTT 2481

Qy   2446 GATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCT 2505
Db   2482 GATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCT 2541

Qy   2506 CTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTT 2565
Db   2542 CTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTT 2601

Qy   2566 TCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATG 2625
Db   2602 TCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATG 2661

Qy   2626 AATGCTACCCAAAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTAT 2685
Db   2662 AATGCTACCCACAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTAT 2721

Qy   2686 AGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGTTTGGTAGT 2745
Db   2722 AGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAGTAAAGGTATTTGGTTTGGTAGT 2781

Qy   2746 GGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTG 2805
Db   2782 GGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTG 2841

Qy   2806 AGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAAT 2865
Db   2842 AGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAAT 2901

Qy   2866 TATCTCAATGGGGAAATTGCTCTGCGATTCTAG 2898
Db   2902 TATCTCAATGGGGAAATTGCTCTGCGATTCTAG 2934
```

RESULT 7
AAH56262
ID AAH56262 standard; DNA; 2934 BP.
XX
AC AAH56262;
XX
DT 05-SEP-2001 (first entry)
XX
DE Chlamydia trachomatis pmpE gene.
XX
KW Chlamydia; vaccine; infection; fusion protein; antigen;
KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;
KW acute respiratory tract infection; Cap1; CT529; OMCB;
KW polymorphic membrane protein; pmp; thiol specific antioxidant; TSA; ds.
XX
OS Chlamydia trachomatis.
XX
PN WO200140474-A2.
XX
PD 07-JUN-2001.
XX
PF 04-DEC-2000; 2000WO-US32919.
XX
PR 03-DEC-1999; 99US-0454684.
PR 19-APR-2000; 2000US-0556877.
PR 20-JUN-2000; 2000US-0598419.
XX
PA (CORI-) CORIXA CORP.
XX
PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Scholler J;
XX
DR WPI; 2001-374831/39.
XX
PT Chlamydia polypeptides and fusion proteins useful for preventing pelvic
PT inflammatory disease, trachoma, acute respiratory tract infections,
PT atherosclerosis and heart disease -
XX
PS Claim 1; Page 207; 295pp; English.
XX
CC The present nucleotide sequence is provided in a specification
CC relating to compounds and methods for the treatment and diagnosis of
CC chlamydial infection. The compounds provided include polypeptides and
CC fusion proteins comprising immunogenic portions of Chlamydia antigens
CC and DNA sequences encoding such polypeptides. They are useful for
CC vaccinating against chlamydial infection, which causes pelvic
CC inflammatory disease, trachoma, acute respiratory tract infections,
CC atherosclerosis and heart disease.
XX
SQ Sequence 2934 BP; 875 A; 645 C; 575 G; 839 T; 0 other;

Query Match 95.9%; Score 2777.8; DB 22; Length 2934;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 2813; Conservative 0; Mismatches 37; Indels 3; Gaps 1;

```
Qy     46 GGACTAGCTAGAGAGGTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCT 105
Db     85 GGACTAGCTAGAGAGGTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCT 144

Qy    106 ACGAAAGAGTCGCTATCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAAC 165
Db    145 ACGAAAGAGTCGCTATCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAAC 204

Qy    166 TGCTATCTCGATAACCTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGA 225
Db    205 TGCTATCTCGATAACCTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGA 264

Qy    226 GCTGCTGTCACAATAACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTAT 285
Db    265 GCTGCTGTCACAATAACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTAT 324

Qy    286 TTTGCAAAAAATCTCACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCT 345
Db    325 TTTGCAAAAAATCTCACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCT 384

Qy    346 CCTACCGTGGAGATTCGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGC 405
Db    385 CCTACCGTGGAGATTCGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGC 444

Qy    406 AGACCATTTACATCGAGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCC 465
Db    445 AGACTATTTACATGGAGAAATCCTTATGCTGCTG---ATAAAATAAGAGAAGGCGGAGCC 501

Qy    466 ATTCATGCTCAAAATCTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAAC 525
Db    502 ATTCATGCTCAAAATCTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAAC 561

Qy    526 TTTTCTTATGTCCGAGGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAAT 585
Db    562 TTTTCTTATGTCCAAGGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAAT 621

Qy    586 CAGTCTTGTTTTCTCTTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGT 645
Db    622 CAGTCTTGTTTTCTCTTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGT 681

Qy    646 GGCGCTATCTATGCTGGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTT 705
Db    682 GGCGCTATCTATGCTGGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTC 741

Qy    706 ATCAATAACGCCTGTTGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGA 765
Db    742 ATCAATAACGCCTGTTGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGA 801

Qy    766 AATCGTGGTAACATCGTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCT 825
Db    802 AATCGTGGTAACATCGTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCT 861

Qy    826 TCAGAAGCTTCTGATGGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAAT 885
Db    862 TCAGAAGCTTCTGATGGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAAT 921
```

```
Qy     886 CGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCT 945
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     922 CGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCT 981

Qy     946 CCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAAT 1005
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     982 CCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAAT 1041

Qy    1006 AAGGGGGGCGCTATCTATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCAT 1065
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    1042 AAGGGGGGCGCTATCTATATAGACGGAACCAGTAACTCCAAAATTTCTGCCGACCGCCAT 1101

Qy    1066 GCTATTATTTTTAATGAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACG 1125
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1102 GCTATTATTTTTAATGAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACG 1161

Qy    1126 TCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTA 1185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1162 TCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTA 1221

Qy    1186 TTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCA 1245
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1222 TTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCA 1281

Qy    1246 GGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGA 1305
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1282 GGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGA 1341

Qy    1306 GCTACTGTTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCC 1365
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1342 GCTACTGTTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCC 1401

Qy    1366 CTTACTCTCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGA 1425
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1402 CTTACTCTCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGA 1461

Qy    1426 TTCACACAAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAA 1485
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1462 TTCACACAAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAA 1521

Qy    1486 AATGGTGCAGGAAATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAAT 1545
           |||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1522 AATGGTACAGGAGATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAAT 1581

Qy    1546 CTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAAT 1605
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1582 CTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAAT 1641

Qy    1606 AACAGCAATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTC 1665
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1642 AACAGCAATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTC 1701

Qy    1666 TCACTCATTGATGACTATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTG 1725
           |||||||||||||||| ||||| ||||||||||||||||||||| ||||||||||||
Db    1702 TCACTCATTGATGACTACGGGAACTCTCCTTATGAATCCACAGATCTGACCCATGCTCTG 1761

Qy    1726 TCATCACAGCCTATGCTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGAT 1785
           |||||||||||||||||||||||||||| ||||| |||||||||||  ||| ||  ||
Db    1762 TCATCACAGCCTATGCTATCTATTTCTGAAGCTAGCGATAACCAGCTACAATCAGAAAAT 1821

Qy    1786 ATGGATTTTTCGGGACTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGC 1845
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1822 ATAGATTTTTCGGGACTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGC 1881

Qy    1846 TGGGCAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAA 1905
           |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    1882 TGGGCAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACTGATCCACAAAAA 1941

Qy    1906 GCCAATAGATTCCATAGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGC 1965
           |||||||||||  |||||||||||  ||||  ||  ||||||||||  ||||||||||||||
Db    1942 GCCAATAGATTTCATAGAACCTTACTACTAACATGGCTTCCTGCCGGGTATGTTCCTAGC 2001

Qy    1966 CCGAAACACAGAAGTCCCCTCATAGCGAATACCTTATGGGGGAATATGCTGCTTGCAACA 2025
           || ||||||||||||||||||||||| || ||||||||||||||||||||||||||||||||
Db    2002 CCAAAACACAGAAGTCCCCTCATAGCTAACACCTTATGGGGGAATATGCTGCTTGCAACA 2061

Qy    2026 GAAAGCTTAAAAAATAGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTACA 2085
           |||||||||||||||||||||| |||||||||||| ||||||||||||||||||||||||
Db    2062 GAAAGCTTAAAAAATAGTGCAGAGCTGACACCTAGTGGTCATCCTTTCTGGGGAATTACA 2121

Qy    2086 GGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCAT 2145
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2122 GGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCAT 2181

Qy    2146 ATGCGCTCTTCCGGATACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTG 2205
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2182 ATGCGCTCTTCCGGATACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTG 2241

Qy    2206 AAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCT 2265
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2242 AAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCT 2301

Qy    2266 AAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACT 2325
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2302 AAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACT 2361

Qy    2326 AAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCACCATTTCTATACCCAAGGA 2385
           |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    2362 AAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCACCATTTCTATACTCAAGGA 2421

Qy    2386 GAAAATCTAACATCTCAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTGTTTTTTTT 2445
           |||||||||||||||||||||||||||||| |||||||||||||||||||| ||||||
Db    2422 GAAAATCTAACATCTCAAGGGACGTTCCGCAGTCAAACGATGGGAGGTGCTGTCTTTTTT 2481

Qy    2446 GATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCT 2505
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2482 GATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCT 2541

Qy    2506 CTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTT 2565
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2542 CTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTT 2601

Qy    2566 TCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATG 2625
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2602 TCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATG 2661

Qy    2626 AATGCTACCCAAAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTAT 2685
           |||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db    2662 AATGCTACCCACAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTAT 2721

Qy    2686 AGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGTTTGGTAGT 2745
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    2722 AGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAGTAAAGGTATTTGGTTTGGTAGT 2781

Qy    2746 GGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTG 2805
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2782 GGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTG 2841

Qy    2806 AGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAAT 2865
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2842 AGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAAT 2901

Qy    2866 TATCTCAATGGGGAAATTGCTCTGCGATTCTAG 2898
           |||||||||||||||||||||||||||||||||
Db    2902 TATCTCAATGGGGAAATTGCTCTGCGATTCTAG 2934
```

```
RESULT 8
ABL92491
ID   ABL92491 standard; DNA; 2934 BP.
XX
AC   ABL92491;
XX
DT   05-JUN-2002  (first entry)
XX
DE   Chlamydia DNA sequence SEQ ID NO:183.
```

XX
KW Chlamydial infection; Chlamydia; vaccine; detection; diagnosis;
KW antigen; antibacterial; immunostimulant; immune response;
KW Chlamydia-specific T-cell response; gene; ds.
XX
OS Chlamydia sp.
XX
PN WO200208267-A2.
XX
PD 31-JAN-2002.
XX
PF 20-JUL-2001; 2001WO-US23121.
XX
PR 20-JUL-2000; 2000US-0620412.
PR 23-APR-2001; 2001US-0841132.
XX
PA (CORI-) CORIXA CORP.
XX
PI Fling SP, Skeiky YAW, Probst P, Bhatia A;
XX
DR WPI; 2002-179901/23.
XX
PT Novel compositions comprising Chlamydia Cap1 protein and its use in the
PT treatment of Chlamydia infection -
XX
PS Example 1; Page 238-239; 537pp; English.
XX
CC The present invention describes compositions comprising a Chlamydia Cap1
CC protein and methods for the diagnosis and therapy of Chlamydia infection.
CC Chlamydia DNA and protein sequences from the present invention can have
CC antibacterial and immunostimulant activities, and can be used in
CC vaccines. Compounds from the present invention can be used for eliciting
CC an immune response, specifically stimulating a Chlamydia-specific T-cell
CC response or inhibiting the development of a Chlamydia infection in an
CC animal. Methods from the present invention can be used: for detecting the
CC presence of Chlamydia in a patient; to stimulate and/or expand T cells
CC specific for a Chlamydia protein; and for treatment of a Chlamydia
CC infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 2934 BP; 875 A; 645 C; 575 G; 839 T; 0 other;

Query Match 95.9%; Score 2777.8; DB 24; Length 2934;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 2813; Conservative 0; Mismatches 37; Indels 3; Gaps 1;

```
Qy     46 GGACTAGCTAGAGAGGTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCT 105
Db     85 GGACTAGCTAGAGAGGTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCT 144

Qy    106 ACGAAAGAGTCGCTATCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAAC 165
Db    145 ACGAAAGAGTCGCTATCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAAC 204

Qy    166 TGCTATCTCGATAACCTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGA 225
Db    205 TGCTATCTCGATAACCTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGA 264

Qy    226 GCTGCTGTCACAATAACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTAT 285
Db    265 GCTGCTGTCACAATAACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTAT 324

Qy    286 TTTGCAAAAAATCTCACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCT 345
Db    325 TTTGCAAAAAATCTCACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCT 384

Qy    346 CCTACCGTGGAGATTCGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGC 405
Db    385 CCTACCGTGGAGATTCGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGC 444

Qy    406 AGACCATTTACATCGAGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCC 465
Db    445 AGACTATTTACATGGAGAAATCCTTATGCTGCTG---ATAAAATAAGAGAAGGCGGAGCC 501

Qy    466 ATTCATGCTCAAAATCTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAAC 525
Db    502 ATTCATGCTCAAAATCTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAAC 561

Qy    526 TTTTCTTATGTCCGAGGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAAT 585
Db    562 TTTTCTTATGTCCAAGGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAAT 621

Qy    586 CAGTCTTGTTTTCTCTTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGT 645
Db    622 CAGTCTTGTTTTCTCTTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGT 681

Qy    646 GGCGCTATCTATGCTGGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTT 705
Db    682 GGCGCTATCTATGCTGGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTC 741

Qy    706 ATCAATAACGCCTGTTGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGA 765
Db    742 ATCAATAACGCCTGTTGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGA 801

Qy    766 AATCGTGGTAACATCGTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCT 825
Db    802 AATCGTGGTAACATCGTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCT 861

Qy    826 TCAGAAGCTTCTGATGGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAAT 885
Db    862 TCAGAAGCTTCTGATGGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAAT 921

Qy    886 CGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCT 945
Db    922 CGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCT 981

Qy    946 CCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAAT 1005
Db    982 CCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAAT 1041

Qy   1006 AAGGGGGGCGCTATCTATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCAT 1065
Db   1042 AAGGGGGGCGCTATCTATATAGACGGAACCAGTAACTCCAAAATTTCTGCCGACCGCCAT 1101

Qy   1066 GCTATTATTTTTAATGAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACG 1125
Db   1102 GCTATTATTTTTAATGAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACG 1161

Qy   1126 TCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTA 1185
Db   1162 TCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTA 1221

Qy   1186 TTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCA 1245
Db   1222 TTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCA 1281

Qy   1246 GGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGA 1305
Db   1282 GGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGA 1341

Qy   1306 GCTACTGTTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCC 1365
Db   1342 GCTACTGTTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCC 1401

Qy   1366 CTTACTCTCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGA 1425
Db   1402 CTTACTCTCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGA 1461

Qy   1426 TTCACACAAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAA 1485
Db   1462 TTCACACAAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAA 1521

Qy   1486 AATGGTGCAGGAAATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAAT 1545
Db   1522 AATGGTACAGGAGATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAAT 1581
```

```
Qy   1546 CTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAAT 1605
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1582 CTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAAT 1641

Qy   1606 AACAGCAATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTC 1665
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1642 AACAGCAATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTC 1701

Qy   1666 TCACTCATTGATGACTATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTG 1725
          |||||||||||||||| ||||| |||||||||||||||||||||| ||||||||||||
Db   1702 TCACTCATTGATGACTACGGGAACTCTCCTTATGAATCCACAGATCTGACCCATGCTCTG 1761

Qy   1726 TCATCACAGCCTATGCTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGAT 1785
          |||||||||||||||||||||||||||| ||||| |||||||||||  ||| ||  ||
Db   1762 TCATCACAGCCTATGCTATCTATTTCTGAAGCTAGCGATAACCAGCTACAATCAGAAAAT 1821

Qy   1786 ATGGATTTTTCGGGACTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGC 1845
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1822 ATAGATTTTTCGGGACTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGC 1881

Qy   1846 TGGGCAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAA 1905
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db   1882 TGGGCAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACTGATCCACAAAAA 1941

Qy   1906 GCCAATAGATTCCATAGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGC 1965
          |||||||||||| |||||||||||| |||| || |||||||||||| ||||||||||||||
Db   1942 GCCAATAGATTTCATAGAACCTTACTACTAACATGGCTTCCTGCCGGGTATGTTCCTAGC 2001

Qy   1966 CCGAAACACAGAAGTCCCCTCATAGCGAATACCTTATGGGGGAATATGCTGCTTGCAACA 2025
          || ||||||||||||||||||||||| || ||||||||||||||||||||||||||||||
Db   2002 CCAAAACACAGAAGTCCCCTCATAGCTAACACCTTATGGGGGAATATGCTGCTTGCAACA 2061

Qy   2026 GAAAGCTTAAAAAAATAGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTACA 2085
          |||||||||||||||||||||| |||||||||||| ||||||||||||||||||||||||
Db   2062 GAAAGCTTAAAAAAATAGTGCAGAGCTGACACCTAGTGGTCATCCTTTCTGGGGAATTACA 2121

Qy   2086 GGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCAT 2145
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2122 GGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCAT 2181

Qy   2146 ATGCGCTCTTCCGGATACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTG 2205
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2182 ATGCGCTCTTCCGGATACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTG 2241

Qy   2206 AAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCT 2265
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2242 AAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCT 2301

Qy   2266 AAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACT 2325
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2302 AAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACT 2361

Qy   2326 AAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCACCATTTCTATACCCAAGGA 2385
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   2362 AAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCACCATTTCTATACTCAAGGA 2421

Qy   2386 GAAAATCTAACATCTCAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTGTTTTTTTT 2445
          |||||||||||||||||||||||||| |||||||||||||||||||||| ||||||
Db   2422 GAAAATCTAACATCTCAAGGGACGTTCCGCAGTCAAACGATGGGAGGTGCTGTCTTTTTT 2481

Qy   2446 GATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCT 2505
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2482 GATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCT 2541

Qy   2506 CTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTT 2565
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2542 CTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTT 2601

Qy   2566 TCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATG 2625
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2602 TCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATG 2661

Qy   2626 AATGCTACCCAAAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTAT 2685
          |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2662 AATGCTACCCACAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTAT 2721

Qy   2686 AGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGTTTGGTAGT 2745
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   2722 AGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAGTAAAGGTATTTGGTTTGGTAGT 2781

Qy   2746 GGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTG 2805
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2782 GGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTG 2841

Qy   2806 AGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAAT 2865
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2842 AGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAAT 2901

Qy   2866 TATCTCAATGGGGAAATTGCTCTGCGATTCTAG 2898
          |||||||||||||||||||||||||||||||
Db   2902 TATCTCAATGGGGAAATTGCTCTGCGATTCTAG 2934
```

RESULT 9
AAD37835
ID AAD37835 standard; DNA; 2871 BP.
XX
AC AAD37835;
XX
DT 27-AUG-2002 (first entry)
XX
DE Chlamydia trachomatis L2 PMPE DNA #2.
XX
KW Polymorphic membrane protein; PMPE; PMPI; vaccine; trachoma; PID; LGV;
KW urethritis; epididymitis; conjunctivitis; cervicitis; cervical cancer;
KW pelvic inflammatory disease; lymphogranuloma venereum; tubal occlusion;
KW endometritis; salpingitis; inflammatory heart disease; cardiomyopathy;
KW infertility; autoimmune myocarditis; atherosclerosis; infection;
KW arthritis; gene; ds.
XX
OS Chlamydia trachomatis.
XX
FH Key Location/Qualifiers
FT CDS 1..2871
FT /*tag= a
FT /product= "Chlamydia trachomatis L2 PMPE protein #2"
XX
PN WO200228998-A2.
XX
PD 11-APR-2002.
XX
PF 28-SEP-2001; 2001WO-US30345.
XX
PR 02-OCT-2000; 2000US-0677752.
XX
PA (ANTE-) ANTEX BIOLOGICS INC.
XX
PI Jackson WJ;
XX
DR WPI; 2002-426107/45.
DR P-PSDB; AAE23486.
XX
PT Novel purified Chlamydia polymorphic membrane protein E or I, useful
PT for preparing vaccines for preventing or treating diseases associated
PT with Chlamydia infection such as trachoma, and infertility -
XX
PS Example; Fig 8; 160pp; English.
XX
CC The invention relates to Chlamydia sp. polymorphic membrane protein
CC (PMP) E and PMPI and nucleic acid molecules encoding such proteins.
CC PMPI and PMPE polypeptides, nucleic acids and vaccines are useful for
CC preventing, treating or ameliorating trachoma, conjunctivitis, tubal
CC occlusion, cervicitis, cervical cancer, pelvic inflammatory disease
CC (PID), lymphogranuloma venereum (LGV), endometritis, epididymitis,
CC salpingitis, infertility, reactive arthritis, inflammatory heart
CC disease, urethritis, autoimmune myocarditis, cardiomyopathy and

CC atherosclerosis. They are also useful as reagents for clinical or
CC medical diagnosis of Chlamydia infections. The present sequence is
CC Chlamydia trachomatis L2 PMPE DNA.
XX
SQ Sequence 2871 BP; 855 A; 623 C; 567 G; 826 T; 0 other;

Query Match 95.8%; Score 2776.8; DB 24; Length 2871;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2787; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```
Qy      94 GTTCCAGATCCTACGAAAGAGTCGCTATCAAATAAAATTAGTTTGACAGGAGACACTCAC 153
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67 GTTCCAGATCCTACGAAAGAGTCGCTATCAAATAAAATTAGTTTGACAGGAGACACTCAC 126

Qy     154 AATCTCACTAACTGCTATCTCGATAACCTACGCTACATACTGGCTATTCTACAAAAAACT 213
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     127 AATCTCACTAACTGCTATCTCGATAACCTACGCTACATACTGGCTATTCTACAAAAAACT 186

Qy     214 CCCAATGAAGGAGCTGCTGTCACAATAACAGATTACCTAAGCTTTTTTGATACACAAAAA 273
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     187 CCCAATGAAGGAGCTGCTGTCACAATAACAGATTACCTAAGCTTTTTTGATACACAAAAA 246

Qy     274 GAAGGTATTTATTTTGCAAAAAATCTCACCCCTGAAAGTGGTGGTGCGATTGGTTATGCG 333
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     247 GAAGGTATTTATTTTGCAAAAAATCTCACCCCTGAAAGTGGTGGTGCGATTGGTTATGCG 306

Qy     334 AGTCCCAATTCTCCTACCGTGGAGATTCGTGATACAATAGGTCCTGTAATCTTTGAAAAT 393
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     307 AGTCCCAATTCTCCTACCGTGGAGATTCGTGATACAATAGGTCCTGTAATCTTTGAAAAT 366

Qy     394 AATACTTGTTGCAGACCATTTACATCGAGTAATCCTAATGCAGCTGTTAATAAAATAAGA 453
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     367 AATACTTGTTGCAGACCATTTACATCGAGTAATCCTAATGCAGCTGTTAATAAAATAAGA 426

Qy     454 GAAGGCGGAGCCATTCATGCTCAAAATCTTTACATAAATCATAATCATGATGTGGTCGGA 513
           ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     427 GAAGGCGGAGCCATTCATGCTCAAAATCTTTACATAAATCACAATCATGATGTGGTCGGA 486

Qy     514 TTTATGAAGAACTTTTCTTATGTCCGAGGAGGAGCCATTAGTACCGCTAATACCTTTGTT 573
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     487 TTTATGAAGAACTTTTCTTATGTCCGAGGAGGAGCCATTAGTACCGCTAATACCTTTGTT 546

Qy     574 GTGAGCGAGAATCAGTCTTGTTTTCTCTTTATGGACAACATCTGTATTCAAACTAATACA 633
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     547 GTGAGCGAGAATCAGTCTTGTTTTCTCTTTATGGACAACATCTGTATTCAAACTAATACA 606

Qy     634 GCAGGAAAAGGTGGCGCTATCTATGCTGGAACGAGCAATTCTTTTGAGAGTAATAACTGC 693
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     607 GCAGGAAAAGGTGGCGCTATCTATGCTGGAACGAGCAATTCTTTTGAGAGTAATAACTGC 666

Qy     694 GATCTCTTCTTTATCAATAACGCCTGTTGTGCAGGAGGAGCGATCTTCTCCCCTATCTGT 753
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     667 GATCTCTTCTTTATCAATAACGCCTGTTGTGCAGGAGGAGCGATCTTCTCCCCTATCTGT 726

Qy     754 TCTCTAACAGGAAATCGTGGTAACATCGTTTTCTATAACAATCGCTGCTTTAAAAATGTA 813
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     727 TCTCTAACAGGAAATCGTGGTAACATCGTTTTCTATAACAATCGCTGCTTTAAAAATGTA 786

Qy     814 GAAACAGCTTCTTCAGAAGCTTCTGATGGAGGAGCAATTAAAGTAACTACTCGCCTAGAT 873
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     787 GAAACAGCTTCTTCAGAAGCTTCTGATGGAGGAGCAATTAAAGTAACTACTCGCCTAGAT 846

Qy     874 GTTACAGGCAATCGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGA 933
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     847 GTTACAGGCAATCGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGA 906

Qy     934 GCTATTTACGCTCCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAACAAT 993
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     907 GCTATTTACGCTCCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAACAAT 966

Qy     994 ATCGCCAATAATAAGGGGGGCGCTATCTATATAGACGGAACCAGCAACTCCAAAATTTCT 1053
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     967 GTCGCCAATAATAAGGGGGGCGCTATCTATATAGACGGAACCAGCAACTCCAAAATTTCT 1026

Qy    1054 GCCGACCGCCATGCTATTATTTTTAATGAAAATATTGTGACTAATGTAACTAATGCAAAT 1113
           |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    1027 GCCGACCGCCATGCTATTATTTTTAATGAAAATATTGTGACTAATGTAACTAGTGCAAAT 1086

Qy    1114 GGTACCAGTACGTCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCT 1173
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1087 GGTACCAGTACGTCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCT 1146

Qy    1174 GGTGAAATTCTATTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAA 1233
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1147 GGTGAAATTCTATTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAA 1206

Qy    1234 GTTAGCAATGCAGGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTA 1293
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1207 GTTAGCAATGCAGGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTA 1266

Qy    1294 GTATTTTCAGGAGCTACTGTTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAA 1353
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1267 GTATTTTCAGGAGCTACTGTTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAA 1326

Qy    1354 ACACCTGCACCCCTTACTCTCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTT 1413
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1327 ACACCTGCACCCCTTACTCTCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTT 1386

Qy    1414 ACAGTGAATCGATTCACACAAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTG 1473
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1387 ACAGTGAATCGATTCACACAAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTG 1446

Qy    1474 AGTTGCTATAAAAATGGTGCAGGAAATTCTGCTAGCAATGCCTCTATAACACTGAAGCAT 1533
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1447 AGTTGCTATAAAAATGGTGCAGGAAATTCTGCTAGCAATGCCTCTATAACACTGAAGCAT 1506

Qy    1534 ATTGGATTGAATCTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTA 1593
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1507 ATTGGATTGAATCTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTA 1566

Qy    1594 GAGCCTACAAATAACAGCAATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGT 1653
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1567 GAGCCTACAAATAACAGCAATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGT 1626

Qy    1654 GATGTAAAACTCTCACTCATTGATGACTATGGGAATTCTCCTTATGAATCCACAGATCTA 1713
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1627 GATGTAAAACTCTCACTCATTGATGACTATGGGAATTCTCCTTATGAATCCACAGATCTA 1686

Qy    1714 ACCCATGCTCTGTCATCACAGCCTATGCTATCTATTTCTGAGGCTAGTGATAACCAGCTA 1773
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1687 ACCCATGCTCTGTCATCACAGCCTATGCTATCTATTTCTGAGGCTAGTGATAACCAGCTA 1746

Qy    1774 AGATCTGATGATATGGATTTTTCGGGACTAAATGTCCCTCATTATGGATGGCAAGGACTT 1833
           |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    1747 AGATCTGATGATATGGATTTCTCGGGACTAAATGTCCCTCATTATGGATGGCAAGGACTT 1806

Qy    1834 TGGACTTGGGGCTGGGCAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACA 1893
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1807 TGGAGTTGGGGCTGGGCAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACA 1866

Qy    1894 GATCCACAAAAAGCCAATAGATTCCATAGAACCTTATTACTGACTTGGCTTCCTGCTGGG 1953
           |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1867 GATCCCAAAAAAGCCAATAGATTCCATAGAACCTTATTACTGACTTGGCTTCCTGCTGGG 1926

Qy    1954 TATGTTCCTAGCCCGAAACACAGAAGTCCCCTCATAGCGAATACCTTATGGGGGAATATG 2013
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1927 TATGTTCCTAGCCCGAAACACAGAAGTCCCCTCATAGCGAATACCTTATGGGGGAATATG 1986

Qy    2014 CTGCTTGCAACAGAAAGCTTAAAAAATAGTGCAGAACTGACACCTAGTGATCATCCTTTC 2073
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1987 CTGCTTGCAACAGAAAGCTTAAAAAATAGTGCAGAACTGACACCTAGTGATCATCCTTTC 2046

Qy    2074 TGGGGAATTACAGGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCAT 2133
           |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    2047 TGGGGAATTACAGGAGGAGGACTAGGCATGATGGTTTACCAAGAACCTCGAGAAAATCAT 2106
```

```
Qy   2134 CCTGGATTCCATATGCGCTCTTCCGGATACTCTGCGGGGATGATAGCAGGGCAGACACAC 2193
          ||||||||||||||||||||||||||||||| |||||||||||||||||||| |||||
Db   2107 CCTGGATTCCATATGCGCTCTTCCGGATACTTTGCGGGGATGATAGCAGGGCAAACACAT 2166

Qy   2194 ACCTTCTCATTGAAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTACGCAAAAAAC 2253
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2167 ACCTTCTCATTGAAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTACGCAAAAAAC 2226

Qy   2254 AACGTATCTTCTAAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAAGGT 2313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2227 AACGTATCTTCTAAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAAGGT 2286

Qy   2314 TTCTTGCTGACTAAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCACCATTTC 2373
          |||||||| |||||||||||||| |||||||||||||||| |||||||||||||||||||
Db   2287 TTCTTGCTGGCTAAATTAGTTGGTCTTTACAGCTATGGAGATCATAACTGTCACCATTTC 2346

Qy   2374 TATACCCAAGGAGAAAATCTAACATCTCAAGGGACGTTCCGTAGTCAAACGATGGGAGGT 2433
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2347 TATACCCAAGGAGAAAATCTAACATCTCAAGGGACGTTCCGTAGTCAAACGATGGGAGGT 2406

Qy   2434 GCTGTTTTTTTTTGATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCTCCC 2493
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2407 GCTGTTTTTTTTTGATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCTCCC 2466

Qy   2494 TTTTTAGGTGCTCTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTGGGAGCCTAT 2553
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2467 TTTTTAGGTGCTCTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTGGGAGCCTAT 2526

Qy   2554 CCGCGAAGCTTTTCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTTAAA 2613
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2527 CCGCGAAGCTTTTCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTTAAA 2586

Qy   2614 GGTAGCTTTATGAATGCTACCCAAAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAA 2673
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2587 GGTAGCTTTATGAATGCTACCCAAAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAA 2646

Qy   2674 CCCGTTCTGTATAGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAGTAAGGGTATT 2733
          |||||||||||||||||||| || |||||||||||||||||||||||||| ||||||
Db   2647 CCCGTTCTGTATAGACAAGAACTAGAGATCGCGACCCAGCTCCTAGCCAGTAAAGGTATT 2706

Qy   2734 TGGTTTGGTAGTGGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAGCAA 2793
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2707 TGGTTTGGTAGTGGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAGCAA 2766

Qy   2794 ACACAACCTTTGAGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCTTCA 2853
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2767 ACACAACCTTTGAGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCTTCA 2826

Qy   2854 ACCTTCTGTAATTATCTCAATGGGGAAATTGCTCTGCGATTCTA 2897
          |||||||||||||||||||||||||||||||||||||||||||
Db   2827 ACCTTCTGTAATTATCTCAATGGGGAAATTGCTCTGCGATTCTA 2870
```

RESULT 10
AAD37817
ID AAD37817 standard; DNA; 1500 BP.
XX
AC AAD37817;
XX
DT 27-AUG-2002 (first entry)
XX
DE Chlamydia trachomatis L2 PMPE DNA fragment #18.
XX
KW Polymorphic membrane protein; PMPE; PMPI; vaccine; trachoma; PID; LGV;
KW urethritis; epididymitis; conjunctivitis; cervicitis; cervical cancer;
KW pelvic inflammatory disease; lymphogranuloma venereum; tubal occlusion;
KW endometritis; salpingitis; inflammatory heart disease; cardiomyopathy;
KW infertility; autoimmune myocarditis; atherosclerosis; infection;
KW arthritis; gene; ds.
XX
OS Chlamydia trachomatis.
XX
FH Key Location/Qualifiers
FT CDS 1..1500
FT /*tag= a
FT /product= "C. trachomatis L2 PMPE protein fragment #18"
FT /note= "CDS does not include stop codon"
FT /partial
XX
PN WO200228998-A2.
XX
PD 11-APR-2002.
XX
PF 28-SEP-2001; 2001WO-US30345.
XX
PR 02-OCT-2000; 2000US-0677752.
XX
PA (ANTE-) ANTEX BIOLOGICS INC.
XX
PI Jackson WJ;
XX
DR WPI; 2002-426107/45.
DR P-PSDB; AAE23472.
XX
PT Novel purified Chlamydia polymorphic membrane protein E or I, useful
PT for preparing vaccines for preventing or treating diseases associated
PT with Chlamydia infection such as trachoma, and infertility -
XX
PS Disclosure; Page 148-149; 160pp; English.
XX
CC The invention relates to Chlamydia sp. polymorphic membrane protein
CC (PMP) E and PMPI and nucleic acid molecules encoding such proteins.
CC PMPI and PMPE polypeptides, nucleic acids and vaccines are useful for
CC preventing, treating or ameliorating trachoma, conjunctivitis, tubal
CC occlusion, cervicitis, cervical cancer, pelvic inflammatory disease
CC (PID), lymphogranuloma venereum (LGV), endometritis, epididymitis,
CC salpingitis, infertility, reactive arthritis, inflammatory heart
CC disease, urethritis, autoimmune myocarditis, cardiomyopathy and
CC atherosclerosis. They are also useful as reagents for clinical or
CC medical diagnosis of Chlamydia infections. The present sequence is
CC Chlamydia trachomatis L2 PMPE DNA fragment.
XX
SQ Sequence 1500 BP; 463 A; 299 C; 287 G; 451 T; 0 other;

Query Match 51.8%; Score 1500; DB 24; Length 1500;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATTGGAAACTCCCTATCAGGACTAGCTAGAGAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATTGGAAACTCCCTATCAGGACTAGCTAGAGAG 60

Qy     61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120

Qy    121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180

Qy    181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240

Qy    241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300

Qy    301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360

Qy    361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACCATTTACATCG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACCATTTACATCG 420
```

```
Qy    421 AGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 480

Qy    481 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCGA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCGA 540

Qy    541 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 600

Qy    601 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 660

Qy    661 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTTATCAATAACGCCTGT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTTATCAATAACGCCTGT 720

Qy    721 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 780

Qy    781 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 840

Qy    841 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 900

Qy    901 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 960

Qy    961 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1020

Qy   1021 TATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1080

Qy   1081 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1140

Qy   1141 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1200

Qy   1201 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 1260

Qy   1261 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 1320

Qy   1321 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 1380

Qy   1381 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 1440

Qy   1441 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTGCAGGAAAT 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTGCAGGAAAT 1500
```

```
RESULT 11
AAS57011/c
ID   AAS57011 standard; DNA; 1634 BP.
XX
AC   AAS57011;
XX
DT   16-JAN-2002  (first entry)
XX
DE   C. trachomatis DNA clone E3-F3-7 encoding partial CT869 and CT870.
XX
KW   Chlamydia; ds; sexually transmitted disease; PID; antibacterial;
KW   pelvic inflammatory disease; antigen; trachoma; gynecological;
KW   acute respiratory tract infection; atherosclerosis; male infertility;
KW   coronary heart disease.
XX
OS   Chlamydia trachomatis.
XX
PN   WO200181379-A2.
XX
PD   01-NOV-2001.
XX
PF   23-APR-2001; 2001WO-US13081.
XX
PR   21-APR-2000; 2000US-198853P.
PR   20-JUL-2000; 2000US-219752P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Bhatia A,  Probst P,  Stromberg EJ;
XX
DR   WPI; 2001-616771/71.
XX
PT   New polynucleotide for treating Chlamydia infections encodes a
PT   polynucleotides containing an immunogenic portion of a Chlamydia
PT   antigen -
XX
PS   Claim 1; Page 130; 208pp; English.
XX
CC   The invention relates to isolated polynucleotide encoding at least
CC   a partial Chlamydia protein which is an antigenic fragment, or the
CC   complements, fragments, homologues and variants, and antibodies
CC   raised against the antigenic proteins (or fragments). The nucleic
CC   acids, proteins and antibodies are used to diagnose and treat Chlamydia
CC   infections (e.g. a sexually transmitted disease, pelvic inflammatory
CC   disease (PID), acute respiratory tract infection, trachoma,
CC   atherosclerosis and coronary heart disease) in a patient, and in
CC   the treatment of male infertility. The compounds of the invention are
CC   also useful for detecting the presence of Chlamydia in a patient, and
CC   stimulating and/or expanding T cells specific for a Chlamydia protein.
CC   The present sequence encodes a Chlamydia antigen.
XX
SQ   Sequence 1634 BP; 485 A; 293 C; 356 G; 500 T; 0 other;

  Query Match             41.1%;  Score 1190.6;  DB 22;  Length 1634;
  Best Local Similarity   99.0%;  Pred. No. 0;
  Matches 1209;  Conservative    0;  Mismatches    9;  Indels    3;  Gaps    1;

Qy      1 ATGAAAAAAAGCGTTTTTCTTTTTCCTTATTGGAAACTCCCTATCAGGACTAGCTAGAGAG 60
          ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   1218 ATGAAAAAAAGCGTTTTTCTTTTTCCTTATCGGAAACTCCCTATCAGGACTAGCTAGAGAG 1159

Qy     61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1158 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 1099

Qy    121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1098 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 1039

Qy    181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1038 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 979
```

```
Qy    241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    978 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 919

Qy    301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    918 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 859

Qy    361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACCATTTACATCG 420
          |||||||||||||||||||||||||||||||||||||||||||||||| |||||||| |
Db    858 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACTATTTACATGG 799

Qy    421 AGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 480
          ||| |||||| |||| ||||   ||||||||||||||||||||||||||||||||||||
Db    798 AGAAATCCTTATGCTGCTG---ATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 742

Qy    481 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCGA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    741 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCAA 682

Qy    541 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    681 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 622

Qy    601 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    621 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 562

Qy    661 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTTATCAATAACGCCTGT 720
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    561 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTCATCAATAACGCCTGT 502

Qy    721 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    501 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 442

Qy    781 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    441 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 382

Qy    841 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    381 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 322

Qy    901 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    321 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 262

Qy    961 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    261 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 202

Qy   1021 TATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1080
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    201 TATATAGACGGAACCAGTAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 142

Qy   1081 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 82

Qy   1141 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 22

Qy   1201 AGCCAAAATTTAATTTTTTAT 1221
          |||||||||||||||||||||
Db     21 AGCCAAAATTTAATTTTTTAT 1

RESULT 12
AAD37801
ID   AAD37801 standard; DNA; 321 BP.
XX
AC   AAD37801;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Chlamydia trachomatis L2 PMPE DNA fragment #2.
XX
KW   Polymorphic membrane protein; PMPE; PMPI; vaccine; trachoma; PID; LGV;
KW   urethritis; epididymitis; conjunctivitis; cervicitis; cervical cancer;
KW   pelvic inflammatory disease; lymphogranuloma venereum; tubal occlusion;
KW   endometritis; salpingitis; inflammatory heart disease; cardiomyopathy;
KW   infertility; autoimmune myocarditis; atherosclerosis; infection;
KW   arthritis; gene; ds.
XX
OS   Chlamydia trachomatis.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..321
FT                   /*tag=  a
FT                   /product= "C. trachomatis L2 PMPE protein fragment #2"
FT                   /note= "CDS does not include start and stop codon"
FT                   /partial
XX
PN   WO200228998-A2.
XX
PD   11-APR-2002.
XX
PF   28-SEP-2001; 2001WO-US30345.
XX
PR   02-OCT-2000; 2000US-0677752.
XX
PA   (ANTE-) ANTEX BIOLOGICS INC.
XX
PI   Jackson WJ;
XX
DR   WPI; 2002-426107/45.
DR   P-PSDB; AAE23456.
XX
PT   Novel purified Chlamydia polymorphic membrane protein E or I, useful
PT   for preparing vaccines for preventing or treating diseases associated
PT   with Chlamydia infection such as trachoma, and infertility -
XX
PS   Disclosure; Page 143; 160pp; English.
XX
CC   The invention relates to Chlamydia sp. polymorphic membrane protein
CC   (PMP) E and PMPI and nucleic acid molecules encoding such proteins.
CC   PMPI and PMPE polypeptides, nucleic acids and vaccines are useful for
CC   preventing, treating or ameliorating trachoma, conjunctivitis, tubal
CC   occlusion, cervicitis, cervical cancer, pelvic inflammatory disease
CC   (PID), lymphogranuloma venereum (LGV), endometritis, epididymitis,
CC   salpingitis, infertility, reactive arthritis, inflammatory heart
CC   disease, urethritis, autoimmune myocarditis, cardiomyopathy and
CC   atherosclerosis. They are also useful as reagents for clinical or
CC   medical diagnosis of Chlamydia infections. The present sequence is
CC   Chlamydia trachomatis L2 PMPE DNA fragment.
XX
SQ   Sequence 321 BP; 101 A; 74 C; 56 G; 90 T; 0 other;

  Query Match             11.1%;  Score 321;  DB 24;  Length 321;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-80;
  Matches  321;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     43 TCAGGACTAGCTAGAGAGGTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGAT 102
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TCAGGACTAGCTAGAGAGGTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGAT 60

Qy    103 CCTACGAAAGAGTCGCTATCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACT 162
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CCTACGAAAGAGTCGCTATCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACT 120

Qy    163 AACTGCTATCTCGATAACCTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAA 222
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AACTGCTATCTCGATAACCTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAA 180
```

```
Qy    223 GGAGCTGCTGTCACAATAACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATT 282
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GGAGCTGCTGTCACAATAACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATT 240

Qy    283 TATTTTGCAAAAAATCTCACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAAT 342
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TATTTTGCAAAAAATCTCACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAAT 300

Qy    343 TCTCCTACCGTGGAGATTCGT 363
          |||||||||||||||||||||
Db    301 TCTCCTACCGTGGAGATTCGT 321


RESULT 13
AAD37805
ID   AAD37805 standard; DNA; 276 BP.
XX
AC   AAD37805;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Chlamydia trachomatis L2 PMPE DNA fragment #6.
XX
KW   Polymorphic membrane protein; PMPE; PMPI; vaccine; trachoma; PID; LGV;
KW   urethritis; epididymitis; conjunctivitis; cervicitis; cervical cancer;
KW   pelvic inflammatory disease; lymphogranuloma venereum; tubal occlusion;
KW   endometritis; salpingitis; inflammatory heart disease; cardiomyopathy;
KW   infertility; autoimmune myocarditis; atherosclerosis; infection;
KW   arthritis; gene; ds.
XX
OS   Chlamydia trachomatis.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..276
FT                   /*tag=  a
FT                   /product= "C. trachomatis L2 PMPE protein fragment #6"
FT                   /note= "CDS does not include start and stop codon"
FT                   /partial
XX
PN   WO200228998-A2.
XX
PD   11-APR-2002.
XX
PF   28-SEP-2001; 2001WO-US30345.
XX
PR   02-OCT-2000; 2000US-0677752.
XX
PA   (ANTE-) ANTEX BIOLOGICS INC.
XX
PI   Jackson WJ;
XX
DR   WPI; 2002-426107/45.
DR   P-PSDB; AAE23460.
XX
PT   Novel purified Chlamydia polymorphic membrane protein E or I, useful
PT   for preparing vaccines for preventing or treating diseases associated
PT   with Chlamydia infection such as trachoma, and infertility -
XX
PS   Disclosure; Page 144; 160pp; English.
XX
CC   The invention relates to Chlamydia sp. polymorphic membrane protein
CC   (PMP) E and PMPI and nucleic acid molecules encoding such proteins.
CC   PMPI and PMPE polypeptides, nucleic acids and vaccines are useful for
CC   preventing, treating or ameliorating trachoma, conjunctivitis, tubal
CC   occlusion, cervicitis, cervical cancer, pelvic inflammatory disease
CC   (PID), lymphogranuloma venereum (LGV), endometritis, epididymitis,
CC   salpingitis, infertility, reactive arthritis, inflammatory heart
CC   disease, urethritis, autoimmune myocarditis, cardiomyopathy and
CC   atherosclerosis. They are also useful as reagents for clinical or
CC   medical diagnosis of Chlamydia infections. The present sequence is
CC   Chlamydia trachomatis L2 PMPE DNA fragment.
XX
SQ   Sequence 276 BP; 87 A; 55 C; 53 G; 81 T; 0 other;

  Query Match             9.5%;  Score 276;  DB 24;  Length 276;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-67;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    823 TCTTCAGAAGCTTCTGATGGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGC 882
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TCTTCAGAAGCTTCTGATGGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGC 60

Qy    883 AATCGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTAC 942
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AATCGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTAC 120

Qy    943 GCTCCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAAT 1002
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCTCCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAAT 180

Qy   1003 AATAAGGGGGGCGCTATCTATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGC 1062
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AATAAGGGGGGCGCTATCTATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGC 240

Qy   1063 CATGCTATTATTTTTAATGAAAATATTGTGACTAAT 1098
          ||||||||||||||||||||||||||||||||||||
Db    241 CATGCTATTATTTTTAATGAAAATATTGTGACTAAT 276


RESULT 14
AAD37802
ID   AAD37802 standard; DNA; 243 BP.
XX
AC   AAD37802;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Chlamydia trachomatis L2 PMPE DNA fragment #3.
XX
KW   Polymorphic membrane protein; PMPE; PMPI; vaccine; trachoma; PID; LGV;
KW   urethritis; epididymitis; conjunctivitis; cervicitis; cervical cancer;
KW   pelvic inflammatory disease; lymphogranuloma venereum; tubal occlusion;
KW   endometritis; salpingitis; inflammatory heart disease; cardiomyopathy;
KW   infertility; autoimmune myocarditis; atherosclerosis; infection;
KW   arthritis; gene; ds.
XX
OS   Chlamydia trachomatis.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..243
FT                   /*tag=  a
FT                   /product= "C. trachomatis L2 PMPE protein fragment #3"
FT                   /note= "CDS does not include start and stop codon"
FT                   /partial
XX
PN   WO200228998-A2.
XX
PD   11-APR-2002.
XX
PF   28-SEP-2001; 2001WO-US30345.
XX
PR   02-OCT-2000; 2000US-0677752.
XX
PA   (ANTE-) ANTEX BIOLOGICS INC.
XX
PI   Jackson WJ;
XX
DR   WPI; 2002-426107/45.
DR   P-PSDB; AAE23457.
XX
PT   Novel purified Chlamydia polymorphic membrane protein E or I, useful
PT   for preparing vaccines for preventing or treating diseases associated
PT   with Chlamydia infection such as trachoma, and infertility -
XX
PS   Disclosure; Page 143; 160pp; English.
XX
```

```
CC   The invention relates to Chlamydia sp. polymorphic membrane protein
CC   (PMP) E and PMPI and nucleic acid molecules encoding such proteins.
CC   PMPI and PMPE polypeptides, nucleic acids and vaccines are useful for
CC   preventing, treating or ameliorating trachoma, conjunctivitis, tubal
CC   occlusion, cervicitis, cervical cancer, pelvic inflammatory disease
CC   (PID), lymphogranuloma venereum (LGV), endometritis, epididymitis,
CC   salpingitis, infertility, reactive arthritis, inflammatory heart
CC   disease, urethritis, autoimmune myocarditis, cardiomyopathy and
CC   atherosclerosis. They are also useful as reagents for clinical or
CC   medical diagnosis of Chlamydia infections. The present sequence is
CC   Chlamydia trachomatis L2 PMPE DNA fragment.
XX
SQ   Sequence 243 BP; 78 A; 54 C; 44 G; 67 T; 0 other;

  Query Match             8.4%;  Score 243;  DB 24;  Length 243;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-58;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     133 AGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAACCTACGCTACATA 192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 AGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAACCTACGCTACATA 60

Qy     193 CTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATAACAGATTACCTA 252
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATAACAGATTACCTA 120

Qy     253 AGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTCACCCCTGAAAGT 312
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTCACCCCTGAAAGT 180

Qy     313 GGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATTCGTGATACAATA 372
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATTCGTGATACAATA 240

Qy     373 GGT 375
           |||
Db     241 GGT 243


RESULT 15
AAD37811
ID   AAD37811 standard; DNA; 228 BP.
XX
AC   AAD37811;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Chlamydia trachomatis L2 PMPE DNA fragment #12.
XX
KW   Polymorphic membrane protein; PMPE; PMPI; vaccine; trachoma; PID; LGV;
KW   urethritis; epididymitis; conjunctivitis; cervicitis; cervical cancer;
KW   pelvic inflammatory disease; lymphogranuloma venereum; tubal occlusion;
KW   endometritis; salpingitis; inflammatory heart disease; cardiomyopathy;
KW   infertility; autoimmune myocarditis; atherosclerosis; infection;
KW   arthritis; gene; ds.
XX
OS   Chlamydia trachomatis.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..228
FT                   /*tag=  a
FT                   /product= "C. trachomatis L2 PMPE protein fragment #12"
FT                   /note= "CDS does not include start and stop codon"
FT                   /partial
XX
PN   WO200228998-A2.
XX
PD   11-APR-2002.
XX
PF   28-SEP-2001; 2001WO-US30345.
XX
PR   02-OCT-2000; 2000US-0677752.
XX
PA   (ANTE-) ANTEX BIOLOGICS INC.
XX
PI   Jackson WJ;
XX
DR   WPI; 2002-426107/45.
DR   P-PSDB; AAE23466.
XX
PT   Novel purified Chlamydia polymorphic membrane protein E or I, useful
PT   for preparing vaccines for preventing or treating diseases associated
PT   with Chlamydia infection such as trachoma, and infertility -
XX
PS   Disclosure; Page 146; 160pp; English.
XX
CC   The invention relates to Chlamydia sp. polymorphic membrane protein
CC   (PMP) E and PMPI and nucleic acid molecules encoding such proteins.
CC   PMPI and PMPE polypeptides, nucleic acids and vaccines are useful for
CC   preventing, treating or ameliorating trachoma, conjunctivitis, tubal
CC   occlusion, cervicitis, cervical cancer, pelvic inflammatory disease
CC   (PID), lymphogranuloma venereum (LGV), endometritis, epididymitis,
CC   salpingitis, infertility, reactive arthritis, inflammatory heart
CC   disease, urethritis, autoimmune myocarditis, cardiomyopathy and
CC   atherosclerosis. They are also useful as reagents for clinical or
CC   medical diagnosis of Chlamydia infections. The present sequence is
CC   Chlamydia trachomatis L2 PMPE DNA fragment.
XX
SQ   Sequence 228 BP; 73 A; 54 C; 48 G; 53 T; 0 other;

  Query Match             7.9%;  Score 228;  DB 24;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-54;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1873 GCATCTTCAGCAACAATCACAGATCCACAAAAAGCCAATAGATTCCATAGAACCTTATTA 1932
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GCATCTTCAGCAACAATCACAGATCCACAAAAAGCCAATAGATTCCATAGAACCTTATTA 60

Qy    1933 CTGACTTGGCTTCCTGCTGGGTATGTTCCTAGCCCGAAACACAGAAGTCCCCTCATAGCG 1992
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CTGACTTGGCTTCCTGCTGGGTATGTTCCTAGCCCGAAACACAGAAGTCCCCTCATAGCG 120

Qy    1993 AATACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAATAGTGCAGAACTG 2052
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AATACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAATAGTGCAGAACTG 180

Qy    2053 ACACCTAGTGATCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 2100
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ACACCTAGTGATCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 228


Search completed: February  4, 2003, 16:35:52
Job time : 2222 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 4, 2003, 15:58:07 ; Search time 110 Seconds
(without alignments)
8079.539 Million cell updates/sec

Title: US-09-677-752-1
Perfect score: 2898
Sequence: 1 atgaaaaaagcgtttttcctt.....aaattgctctgcgattctag 2898

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2821.2	97.3	2895	4	US-09-556-877-171	Sequence 171, App
	2	2821.2	97.3	2895	4	US-09-620-412C-171	Sequence 171, App
	3	2777.8	95.9	2934	4	US-09-556-877-183	Sequence 183, App
	4	2777.8	95.9	2934	4	US-09-620-412C-183	Sequence 183, App
	5	48.4	1.7	2949	4	US-09-556-877-170	Sequence 170, App
	6	48.4	1.7	2949	4	US-09-620-412C-170	Sequence 170, App
	7	48.4	1.7	3021	4	US-09-556-877-182	Sequence 182, App
	8	48.4	1.7	3021	4	US-09-620-412C-182	Sequence 182, App
	9	45.2	1.6	1941	4	US-09-620-412C-316	Sequence 316, App
	10	41.4	1.4	1896	4	US-09-620-412C-324	Sequence 324, App
c	11	40.6	1.4	1122	2	US-08-915-107-3	Sequence 3, Appli
c	12	40.6	1.4	1122	4	US-09-273-613-3	Sequence 3, Appli
c	13	40.6	1.4	1122	4	US-09-310-867-3	Sequence 3, Appli
c	14	40.6	1.4	1125	2	US-08-915-107-1	Sequence 1, Appli
c	15	40.6	1.4	1125	4	US-09-273-613-1	Sequence 1, Appli
c	16	40.6	1.4	1125	4	US-09-310-867-1	Sequence 1, Appli
	17	40.6	1.4	1860	4	US-09-620-412C-308	Sequence 308, App
	18	39.2	1.4	1560	1	US-08-356-180-1	Sequence 1, Appli
	19	38.8	1.3	1557	4	US-09-620-412C-332	Sequence 332, App
	20	38.8	1.3	5265	4	US-09-556-877-174	Sequence 174, App
	21	38.8	1.3	5265	4	US-09-620-412C-174	Sequence 174, App
	22	38.4	1.3	807	3	US-08-718-905-1	Sequence 1, Appli
	23	38.4	1.3	807	4	US-09-550-497-1	Sequence 1, Appli
	24	38.4	1.3	807	4	US-09-147-992-1	Sequence 1, Appli
	25	38.4	1.3	807	4	US-09-549-839A-1	Sequence 1, Appli
	26	38.4	1.3	1590	4	US-09-549-839A-11	Sequence 11, Appl
	27	38.2	1.3	1011	4	US-09-134-001C-2460	Sequence 2460, Ap
	28	38	1.3	2547	4	US-09-556-877-184	Sequence 184, App
	29	38	1.3	2547	4	US-09-620-412C-184	Sequence 184, App
	30	38	1.3	4593	4	US-09-556-877-172	Sequence 172, App
	31	38	1.3	4593	4	US-09-620-412C-172	Sequence 172, App
c	32	36.4	1.3	1263	4	US-09-443-184-37	Sequence 37, Appl
	33	36.2	1.2	1311	4	US-08-584-760A-66	Sequence 66, Appl
	34	36	1.2	8703	4	US-08-961-527-177	Sequence 177, App
	35	35.4	1.2	3513	1	US-07-828-788A-5	Sequence 5, Appli
	36	35.4	1.2	3513	1	US-08-278-685-1	Sequence 1, Appli
	37	35.4	1.2	3513	1	US-08-278-685-3	Sequence 3, Appli
	38	35.4	1.2	3513	1	US-08-277-721-1	Sequence 1, Appli
	39	35.4	1.2	3513	1	US-08-277-721-3	Sequence 3, Appli
	40	35.4	1.2	3513	1	US-08-602-737-3	Sequence 3, Appli
	41	35.4	1.2	3513	4	US-09-001-982-3	Sequence 3, Appli
	42	35.4	1.2	3513	5	PCT-US92-11337-5	Sequence 5, Appli
c	43	35	1.2	921	4	US-09-247-155-133	Sequence 133, App
	44	35	1.2	2148	4	US-09-620-412C-320	Sequence 320, App
	45	35	1.2	6909	2	US-08-804-196-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-556-877-171
; Sequence 171, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 171
; LENGTH: 2895
; TYPE: DNA
; ORGANISM: Chlamydia
US-09-556-877-171

Query Match 97.3%; Score 2821.2; DB 4; Length 2895;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 2857; Conservative 0; Mismatches 38; Indels 3; Gaps 1;

```
Qy       1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATTGGAAACTCCCTATCAGGACTAGCTAGAGAG 60
           |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db       1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATCGGAAACTCCCTATCAGGACTAGCTAGAGAG 60

Qy      61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120

Qy     121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180

Qy     181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240

Qy     241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300

Qy     301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360
```

```
Qy    361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACCATTTACATCG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||| |
Db    361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACTATTTACATGG 420

Qy    421 AGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 480
          || |||||| |||| |||| ||||||||||||||||||||||||||||||||||||||||
Db    421 AGAAATCCTTATGCTGCTG---ATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 477

Qy    481 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCGA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    478 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCAA 537

Qy    541 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    538 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 597

Qy    601 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    598 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 657

Qy    661 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTTATCAATAACGCCTGT 720
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    658 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTCATCAATAACGCCTGT 717

Qy    721 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    718 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 777

Qy    781 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    778 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 837

Qy    841 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    838 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 897

Qy    901 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    898 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 957

Qy    961 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    958 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1017

Qy   1021 TATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1080
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   1018 TATATAGACGGAACCAGTAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1077

Qy   1081 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1078 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1137

Qy   1141 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1138 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1197

Qy   1201 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1198 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 1257

Qy   1261 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1258 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 1317

Qy   1321 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1318 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 1377

Qy   1381 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1378 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 1437

Qy   1441 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTGCAGGAAAT 1500
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||| ||
Db   1438 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTACAGGAGAT 1497

Qy   1501 TCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTG 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1498 TCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTG 1557

Qy   1561 AAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTAT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1558 AAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTAT 1617

Qy   1621 ACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGAC 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1618 ACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGAC 1677

Qy   1681 TATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTGTCATCACAGCCTATG 1740
          || ||||| |||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   1678 TACGGGAACTCTCCTTATGAATCCACAGATCTGACCCATGCTCTGTCATCACAGCCTATG 1737

Qy   1741 CTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGATATGGATTTTTCGGGA 1800
          |||||||||||||| ||||| |||||||||||  ||| ||  |||| ||||||||||||
Db   1738 CTATCTATTTCTGAAGCTAGCGATAACCAGCTACAATCAGAAAATATAGATTTTTCGGGA 1797

Qy   1801 CTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1798 CTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAA 1857

Qy   1861 GATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAAGCCAATAGATTCCAT 1920
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||| |||
Db   1858 GATCCAGAACCAGCATCTTCAGCAACAATCACTGATCCACAAAAAGCCAATAGATTTCAT 1917

Qy   1921 AGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGCCCGAAACACAGAAGT 1980
          |||||||| |||| || |||||||||| |||||||||||||||| ||||||||||||
Db   1918 AGAACCTTACTACTAACATGGCTTCCTGCCGGGTATGTTCCTAGCCCAAAACACAGAAGT 1977

Qy   1981 CCCCTCATAGCGAATACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAAT 2040
          |||||||||| || ||||||||||||||||||||||||||||||||||||||||||||||
Db   1978 CCCCTCATAGCTAACACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAAT 2037

Qy   2041 AGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 2100
          |||||||| |||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   2038 AGTGCAGAGCTGACACCTAGTGGTCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 2097

Qy   2101 ATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGA 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2098 ATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGA 2157

Qy   2161 TACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTGAAATTCAGTCAGACC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2158 TACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTGAAATTCAGTCAGACC 2217

Qy   2221 TACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2218 TACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGC 2277

Qy   2281 CAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2278 CAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTT 2337

Qy   2341 TACAGCTATGGAGACCATAACTGTCACCATTTCTATACCCAAGGAGAAAATCTAACATCT 2400
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   2338 TACAGCTATGGAGACCATAACTGTCACCATTTCTATACTCAAGGAGAAAATCTAACATCT 2397

Qy   2401 CAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTGTTTTTTTTGATCTCCCTATGAAA 2460
          |||||||||||||| |||||||||||||||||||||||| ||||||||||||||||||||
Db   2398 CAAGGGACGTTCCGCAGTCAAACGATGGGAGGTGCTGTCTTTTTTGATCTCCCTATGAAA 2457

Qy   2461 CCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCT 2520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2458 CCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCT 2517

Qy   2521 AGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCT 2580
```

```
Db    2518 AGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCT 2577

Qy    2581 TTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCAAAGA 2640
Db    2578 TTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCACAGA 2637

Qy    2641 CCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACCAGGG 2700
Db    2638 CCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACCAGGG 2697

Qy    2701 ATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCG 2760
Db    2698 ATCGCGACCCAGCTCCTAGCCAGTAAAGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCG 2757

Qy    2761 CGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTC 2820
Db    2758 CGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTC 2817

Qy    2821 CATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAA 2880
Db    2818 CATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAA 2877

Qy    2881 ATTGCTCTGCGATTCTAG 2898
Db    2878 ATTGCTCTGCGATTCTAG 2895


RESULT 2
US-09-620-412C-171
; Sequence 171, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 171
;   LENGTH: 2895
;   TYPE: DNA
;   ORGANISM: Chlamydia
US-09-620-412C-171

  Query Match             97.3%;  Score 2821.2;  DB 4;  Length 2895;
  Best Local Similarity   98.6%;  Pred. No. 0;
  Matches 2857;  Conservative    0;  Mismatches   38;  Indels    3;  Gaps    1;

Qy       1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATTGGAAACTCCCTATCAGGACTAGCTAGAGAG 60
Db       1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATCGGAAACTCCCTATCAGGACTAGCTAGAGAG 60

Qy      61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120
Db      61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120

Qy     121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180
Db     121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180

Qy     181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240
Db     181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240

Qy     241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300
Db     241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300

Qy     301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360
Db     301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360

Qy     361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACCATTTACATCG 420
Db     361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACTATTTACATGG 420

Qy     421 AGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 480
Db     421 AGAAATCCTTATGCTGCTG---ATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 477

Qy     481 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCGA 540
Db     478 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCAA 537

Qy     541 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 600
Db     538 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 597

Qy     601 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 660
Db     598 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 657

Qy     661 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTTATCAATAACGCCTGT 720
Db     658 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTCATCAATAACGCCTGT 717

Qy     721 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 780
Db     718 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 777

Qy     781 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 840
Db     778 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 837

Qy     841 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 900
Db     838 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 897

Qy     901 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 960
Db     898 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 957

Qy     961 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1020
Db     958 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1017

Qy    1021 TATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1080
Db    1018 TATATAGACGGAACCAGTAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1077

Qy    1081 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1140
Db    1078 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1137

Qy    1141 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1200
Db    1138 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1197

Qy    1201 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 1260
Db    1198 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 1257

Qy    1261 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 1320
Db    1258 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 1317

Qy    1321 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 1380
Db    1318 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 1377

Qy    1381 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 1440
Db    1378 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 1437
```

```
Qy   1441 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTGCAGGAAAT 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||| ||
Db   1438 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTACAGGAGAT 1497

Qy   1501 TCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTG 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1498 TCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTG 1557

Qy   1561 AAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTAT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1558 AAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTAT 1617

Qy   1621 ACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGAC 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1618 ACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGAC 1677

Qy   1681 TATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTGTCATCACAGCCTATG 1740
          || |||||| ||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   1678 TACGGGAACTCTCCTTATGAATCCACAGATCTGACCCATGCTCTGTCATCACAGCCTATG 1737

Qy   1741 CTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGATATGGATTTTTCGGGA 1800
          |||||||||||||| |||| |||||||||||  ||| ||  |||| ||||||||||||
Db   1738 CTATCTATTTCTGAAGCTAGCGATAACCAGCTACAATCAGAAAATATAGATTTTTCGGGA 1797

Qy   1801 CTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1798 CTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAA 1857

Qy   1861 GATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAAGCCAATAGATTCCAT 1920
          |||||||||||||||||||||||||||||| |||||||||||||||||||||| |||
Db   1858 GATCCAGAACCAGCATCTTCAGCAACAATCACTGATCCACAAAAAGCCAATAGATTTCAT 1917

Qy   1921 AGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGCCCGAAACACAGAAGT 1980
          ||||||||| |||| || |||||||||||| ||||||||||||||||| ||||||||||
Db   1918 AGAACCTTACTACTAACATGGCTTCCTGCCGGGTATGTTCCTAGCCCAAAACACAGAAGT 1977

Qy   1981 CCCCTCATAGCGAATACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAAT 2040
          |||||||||| || ||||||||||||||||||||||||||||||||||||||||||||||
Db   1978 CCCCTCATAGCTAACACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAAT 2037

Qy   2041 AGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 2100
          |||||||| |||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   2038 AGTGCAGAGCTGACACCTAGTGGTCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 2097

Qy   2101 ATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGA 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2098 ATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGA 2157

Qy   2161 TACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTGAAATTCAGTCAGACC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2158 TACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTGAAATTCAGTCAGACC 2217

Qy   2221 TACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2218 TACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGC 2277

Qy   2281 CAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2278 CAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTT 2337

Qy   2341 TACAGCTATGGAGACCATAACTGTCACCATTTCTATACCCAAGGAGAAAATCTAACATCT 2400
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   2338 TACAGCTATGGAGACCATAACTGTCACCATTTCTATACTCAAGGAGAAAATCTAACATCT 2397

Qy   2401 CAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTGTTTTTTTTGATCTCCCTATGAAA 2460
          |||||||||||||| |||||||||||||||||||||| ||||||||||||||||||||||
Db   2398 CAAGGGACGTTCCGCAGTCAAACGATGGGAGGTGCTGTCTTTTTTGATCTCCCTATGAAA 2457

Qy   2461 CCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2458 CCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCT 2517
```

```
Qy   2521 AGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCT 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2518 AGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCT 2577

Qy   2581 TTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCAAAGA 2640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   2578 TTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCACAGA 2637

Qy   2641 CCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACCAGGG 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2638 CCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACCAGGG 2697

Qy   2701 ATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCG 2760
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   2698 ATCGCGACCCAGCTCCTAGCCAGTAAAGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCG 2757

Qy   2761 CGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTC 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2758 CGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTC 2817

Qy   2821 CATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAA 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2818 CATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAA 2877

Qy   2881 ATTGCTCTGCGATTCTAG 2898
          ||||||||||||||||||
Db   2878 ATTGCTCTGCGATTCTAG 2895
```

RESULT 3
US-09-556-877-183
; Sequence 183, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 183
; LENGTH: 2934
; TYPE: DNA
; ORGANISM: Chlamydia
US-09-556-877-183

Query Match 95.9%; Score 2777.8; DB 4; Length 2934;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 2813; Conservative 0; Mismatches 37; Indels 3; Gaps 1;

```
Qy     46 GGACTAGCTAGAGAGGTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCT 105
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     85 GGACTAGCTAGAGAGGTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCT 144

Qy    106 ACGAAAGAGTCGCTATCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAAC 165
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    145 ACGAAAGAGTCGCTATCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAAC 204

Qy    166 TGCTATCTCGATAACCTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGA 225
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    205 TGCTATCTCGATAACCTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGA 264

Qy    226 GCTGCTGTCACAATAACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTAT 285
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    265 GCTGCTGTCACAATAACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTAT 324

Qy    286 TTTGCAAAAAATCTCACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCT 345
```


Db 325 TTTGCAAAAAATCTCACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCT 384
Qy 346 CCTACCGTGGAGATTCGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGC 405
Db 385 CCTACCGTGGAGATTCGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGC 444
Qy 406 AGACCATTTACATCGAGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCC 465
Db 445 AGACTATTTACATGGAGAAATCCTTATGCTGCTG---ATAAAATAAGAGAAGGCGGAGCC 501
Qy 466 ATTCATGCTCAAAATCTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAAC 525
Db 502 ATTCATGCTCAAAATCTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAAC 561
Qy 526 TTTTCTTATGTCCGAGGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAAT 585
Db 562 TTTTCTTATGTCCAAGGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAAT 621
Qy 586 CAGTCTTGTTTTCTCTTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGT 645
Db 622 CAGTCTTGTTTTCTCTTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGT 681
Qy 646 GGCGCTATCTATGCTGGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTT 705
Db 682 GGCGCTATCTATGCTGGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTC 741
Qy 706 ATCAATAACGCCTGTTGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGA 765
Db 742 ATCAATAACGCCTGTTGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGA 801
Qy 766 AATCGTGGTAACATCGTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCT 825
Db 802 AATCGTGGTAACATCGTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCT 861
Qy 826 TCAGAAGCTTCTGATGGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAAT 885
Db 862 TCAGAAGCTTCTGATGGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAAT 921
Qy 886 CGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCT 945
Db 922 CGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCT 981
Qy 946 CCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAAT 1005
Db 982 CCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAAT 1041
Qy 1006 AAGGGGGGCGCTATCTATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCAT 1065
Db 1042 AAGGGGGGCGCTATCTATATAGACGGAACCAGTAACTCCAAAATTTCTGCCGACCGCCAT 1101
Qy 1066 GCTATTATTTTTAATGAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACG 1125
Db 1102 GCTATTATTTTTAATGAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACG 1161
Qy 1126 TCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTA 1185
Db 1162 TCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTA 1221
Qy 1186 TTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCA 1245
Db 1222 TTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCA 1281
Qy 1246 GGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGA 1305
Db 1282 GGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGA 1341
Qy 1306 GCTACTGTTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCC 1365
Db 1342 GCTACTGTTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCC 1401
Qy 1366 CTTACTCTCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGA 1425
Db 1402 CTTACTCTCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGA 1461
Qy 1426 TTCACACAAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAA 1485
Db 1462 TTCACACAAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAA 1521
Qy 1486 AATGGTGCAGGAAATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAAT 1545
Db 1522 AATGGTACAGGAGATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAAT 1581
Qy 1546 CTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAAT 1605
Db 1582 CTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAAT 1641
Qy 1606 AACAGCAATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTC 1665
Db 1642 AACAGCAATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTC 1701
Qy 1666 TCACTCATTGATGACTATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTG 1725
Db 1702 TCACTCATTGATGACTACGGGAACTCTCCTTATGAATCCACAGATCTGACCCATGCTCTG 1761
Qy 1726 TCATCACAGCCTATGCTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGAT 1785
Db 1762 TCATCACAGCCTATGCTATCTATTTCTGAAGCTAGCGATAACCAGCTACAATCAGAAAAT 1821
Qy 1786 ATGGATTTTTCGGGACTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGC 1845
Db 1822 ATAGATTTTTCGGGACTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGC 1881
Qy 1846 TGGGCAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAA 1905
Db 1882 TGGGCAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACTGATCCACAAAAA 1941
Qy 1906 GCCAATAGATTCCATAGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGC 1965
Db 1942 GCCAATAGATTTCATAGAACCTTACTACTAACATGGCTTCCTGCCGGGTATGTTCCTAGC 2001
Qy 1966 CCGAAACACAGAAGTCCCCTCATAGCGAATACCTTATGGGGGAATATGCTGCTTGCAACA 2025
Db 2002 CCAAAACACAGAAGTCCCCTCATAGCTAACACCTTATGGGGGAATATGCTGCTTGCAACA 2061
Qy 2026 GAAAGCTTAAAAAATAGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTACA 2085
Db 2062 GAAAGCTTAAAAAATAGTGCAGAGCTGACACCTAGTGGTCATCCTTTCTGGGGAATTACA 2121
Qy 2086 GGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCAT 2145
Db 2122 GGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCAT 2181
Qy 2146 ATGCGCTCTTCCGGATACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTG 2205
Db 2182 ATGCGCTCTTCCGGATACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTG 2241
Qy 2206 AAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCT 2265
Db 2242 AAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCT 2301
Qy 2266 AAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACT 2325
Db 2302 AAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACT 2361
Qy 2326 AAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCACCATTTCTATACCCAAGGA 2385
Db 2362 AAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCACCATTTCTATACTCAAGGA 2421
Qy 2386 GAAAATCTAACATCTCAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTGTTTTTTTT 2445
Db 2422 GAAAATCTAACATCTCAAGGGACGTTCCGCAGTCAAACGATGGGAGGTGCTGTCTTTTTT 2481
Qy 2446 GATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCT 2505
Db 2482 GATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCT 2541

```
Qy   2506 CTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTT 2565
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2542 CTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTT 2601

Qy   2566 TCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATG 2625
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2602 TCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATG 2661

Qy   2626 AATGCTACCCAAAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTAT 2685
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2662 AATGCTACCCACAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTAT 2721

Qy   2686 AGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGTTTGGTAGT 2745
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   2722 AGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAGTAAAGGTATTTGGTTTGGTAGT 2781

Qy   2746 GGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTG 2805
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2782 GGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTG 2841

Qy   2806 AGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAAT 2865
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2842 AGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAAT 2901

Qy   2866 TATCTCAATGGGGAAATTGCTCTGCGATTCTAG 2898
          ||||||||||||||||||||||||||||||||
Db   2902 TATCTCAATGGGGAAATTGCTCTGCGATTCTAG 2934


RESULT 4
US-09-620-412C-183
; Sequence 183, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 183
;   LENGTH: 2934
;   TYPE: DNA
;   ORGANISM: Chlamydia
US-09-620-412C-183

  Query Match             95.9%;  Score 2777.8;  DB 4;  Length 2934;
  Best Local Similarity   98.6%;  Pred. No. 0;
  Matches 2813;  Conservative    0;  Mismatches   37;  Indels    3;  Gaps    1;

Qy     46 GGACTAGCTAGAGAGGTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCT 105
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     85 GGACTAGCTAGAGAGGTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCT 144

Qy    106 ACGAAAGAGTCGCTATCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAAC 165
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    145 ACGAAAGAGTCGCTATCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAAC 204

Qy    166 TGCTATCTCGATAACCTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGA 225
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    205 TGCTATCTCGATAACCTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGA 264

Qy    226 GCTGCTGTCACAATAACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTAT 285
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    265 GCTGCTGTCACAATAACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTAT 324

Qy    286 TTTGCAAAAAATCTCACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCT 345
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    325 TTTGCAAAAAATCTCACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCT 384

Qy    346 CCTACCGTGGAGATTCGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGC 405
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    385 CCTACCGTGGAGATTCGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGC 444

Qy    406 AGACCATTTACATCGAGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCC 465
          |||| |||||||| ||| |||||| |||| ||||   ||||||||||||||||||||||
Db    445 AGACTATTTACATGGAGAAATCCTTATGCTGCTG---ATAAAATAAGAGAAGGCGGAGCC 501

Qy    466 ATTCATGCTCAAAATCTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAAC 525
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    502 ATTCATGCTCAAAATCTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAAC 561

Qy    526 TTTTCTTATGTCCGAGGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAAT 585
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    562 TTTTCTTATGTCCAAGGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAAT 621

Qy    586 CAGTCTTGTTTTCTCTTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGT 645
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    622 CAGTCTTGTTTTCTCTTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGT 681

Qy    646 GGCGCTATCTATGCTGGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTT 705
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    682 GGCGCTATCTATGCTGGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTC 741

Qy    706 ATCAATAACGCCTGTTGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGA 765
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    742 ATCAATAACGCCTGTTGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGA 801

Qy    766 AATCGTGGTAACATCGTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCT 825
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    802 AATCGTGGTAACATCGTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCT 861

Qy    826 TCAGAAGCTTCTGATGGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAAT 885
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    862 TCAGAAGCTTCTGATGGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAAT 921

Qy    886 CGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCT 945
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    922 CGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCT 981

Qy    946 CCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAAT 1005
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    982 CCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAAT 1041

Qy   1006 AAGGGGGGCGCTATCTATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCAT 1065
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   1042 AAGGGGGGCGCTATCTATATAGACGGAACCAGTAACTCCAAAATTTCTGCCGACCGCCAT 1101

Qy   1066 GCTATTATTTTTAATGAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACG 1125
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1102 GCTATTATTTTTAATGAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACG 1161

Qy   1126 TCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTA 1185
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1162 TCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTA 1221

Qy   1186 TTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCA 1245
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1222 TTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCA 1281

Qy   1246 GGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGA 1305
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1282 GGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGA 1341

Qy   1306 GCTACTGTTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCC 1365
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1342 GCTACTGTTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCC 1401

Qy   1366 CTTACTCTCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGA 1425
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1402 CTTACTCTCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGA 1461

Qy   1426 TTCACACAAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAA 1485
```

```
Db   1462 TTCACACAAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAA 1521

Qy   1486 AATGGTGCAGGAAATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAAT 1545

Db   1522 AATGGTACAGGAGATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAAT 1581

Qy   1546 CTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAAT 1605

Db   1582 CTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAAT 1641

Qy   1606 AACAGCAATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTC 1665

Db   1642 AACAGCAATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTC 1701

Qy   1666 TCACTCATTGATGACTATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTG 1725

Db   1702 TCACTCATTGATGACTACGGGAACTCTCCTTATGAATCCACAGATCTGACCCATGCTCTG 1761

Qy   1726 TCATCACAGCCTATGCTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGAT 1785

Db   1762 TCATCACAGCCTATGCTATCTATTTCTGAAGCTAGCGATAACCAGCTACAATCAGAAAAT 1821

Qy   1786 ATGGATTTTTCGGGACTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGC 1845

Db   1822 ATAGATTTTTCGGGACTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGC 1881

Qy   1846 TGGGCAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAA 1905

Db   1882 TGGGCAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACTGATCCACAAAAA 1941

Qy   1906 GCCAATAGATTCCATAGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGC 1965

Db   1942 GCCAATAGATTTCATAGAACCTTACTACTAACATGGCTTCCTGCCGGGTATGTTCCTAGC 2001

Qy   1966 CCGAAACACAGAAGTCCCCTCATAGCGAATACCTTATGGGGGAATATGCTGCTTGCAACA 2025

Db   2002 CCAAAACACAGAAGTCCCCTCATAGCTAACACCTTATGGGGGAATATGCTGCTTGCAACA 2061

Qy   2026 GAAAGCTTAAAAAATAGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTACA 2085

Db   2062 GAAAGCTTAAAAAATAGTGCAGAGCTGACACCTAGTGGTCATCCTTTCTGGGGAATTACA 2121

Qy   2086 GGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCAT 2145

Db   2122 GGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCAT 2181

Qy   2146 ATGCGCTCTTCCGGATACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTG 2205

Db   2182 ATGCGCTCTTCCGGATACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTG 2241

Qy   2206 AAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCT 2265

Db   2242 AAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCT 2301

Qy   2266 AAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACT 2325

Db   2302 AAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACT 2361

Qy   2326 AAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCACCATTTCTATACCCAAGGA 2385

Db   2362 AAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCACCATTTCTATACTCAAGGA 2421

Qy   2386 GAAAATCTAACATCTCAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTGTTTTTTTT 2445

Db   2422 GAAAATCTAACATCTCAAGGGACGTTCCGCAGTCAAACGATGGGAGGTGCTGTCTTTTTT 2481

Qy   2446 GATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCT 2505

Db   2482 GATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCT 2541

Qy   2506 CTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTT 2565

Db   2542 CTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTT 2601

Qy   2566 TCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATG 2625

Db   2602 TCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATG 2661

Qy   2626 AATGCTACCCAAAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTAT 2685

Db   2662 AATGCTACCCACAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTAT 2721

Qy   2686 AGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGTTTGGTAGT 2745

Db   2722 AGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAGTAAAGGTATTTGGTTTGGTAGT 2781

Qy   2746 GGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTG 2805

Db   2782 GGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTG 2841

Qy   2806 AGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAAT 2865

Db   2842 AGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAAT 2901

Qy   2866 TATCTCAATGGGGAAATTGCTCTGCGATTCTAG 2898

Db   2902 TATCTCAATGGGGAAATTGCTCTGCGATTCTAG 2934


RESULT 5
US-09-556-877-170
; Sequence 170, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
;  APPLICANT:  Probst, Peter
;  APPLICANT:  Bhatia, Ajay
;  APPLICANT:  Skeiky, Yasir
;  APPLICANT:  Fling, Steve
;  APPLICANT:  Maisonneuve, Jeff
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE:  210121.469C5
;  CURRENT APPLICATION NUMBER: US/09/556,877
;  CURRENT FILING DATE:  2000-04-19
;  NUMBER OF SEQ ID NOS:  305
;  SOFTWARE:  FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 170
;   LENGTH: 2949
;   TYPE: DNA
;   ORGANISM: Chlamydia
US-09-556-877-170

  Query Match             1.7%;  Score 48.4;  DB 4;  Length 2949;
  Best Local Similarity   56.7%;  Pred. No. 0.001;
  Matches  110;  Conservative    0;  Mismatches   81;  Indels    3;  Gaps    1;

Qy    896 TCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTA 955

Db   1001 TTTTCTTTAGTAGCAATGTAGCTGCTGGGAAAGGGGGAGCTATTTATGCCAAAAAGCTCT 1060

Qy    956 CCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCG 1015

Db   1061 CGGTTGCTAACTGTGGCCCTGTACAATTTTTAAGGAATATCGCTAATGAT---GGTGGAG 1117

Qy   1016 CTATCTATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTT 1075

Db   1118 CGATTTATTTAGGAGAATCTGGAGAGCTCAGTTTATCTGCTGATTATGGAGATATTATTT 1177

Qy   1076 TTAATGAAAATATT 1089

Db   1178 TCGATGGGAATCTT 1191


RESULT 6
US-09-620-412C-170
```

```
; Sequence 170, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 170
;   LENGTH: 2949
;   TYPE: DNA
;   ORGANISM: Chlamydia
US-09-620-412C-170

  Query Match             1.7%;  Score 48.4;  DB 4;  Length 2949;
  Best Local Similarity   56.7%;  Pred. No. 0.001;
  Matches  110;  Conservative    0;  Mismatches   81;  Indels    3;  Gaps    1;

Qy     896 TCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTA 955
           | || ||||||  |||| |  |         | || ||||||||||| ||        |
Db    1001 TTTTCTTTAGTAGCAATGTAGCTGCTGGGAAAGGGGGAGCTATTTATGCCAAAAAGCTCT 1060

Qy     956 CCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCG 1015
           |  | |   |  |||||||    | ||| |||  |||||||| ||| ||   || || |
Db    1061 CGGTTGCTAACTGTGGCCCTGTACAATTTTTAAGGAATATCGCTAATGAT---GGTGGAG 1117

Qy    1016 CTATCTATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTT 1075
           | || ||| |||  | | |  |  |   ||   | ||||| ||        | ||||||||
Db    1118 CGATTTATTTAGGAGAATCTGGAGAGCTCAGTTTATCTGCTGATTATGGAGATATTATTT 1177

Qy    1076 TTAATGAAAATATT 1089
           |  |||  ||| ||
Db    1178 TCGATGGGAATCTT 1191


RESULT 7
US-09-556-877-182
; Sequence 182, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
;  APPLICANT:  Probst, Peter
;  APPLICANT:  Bhatia, Ajay
;  APPLICANT:  Skeiky, Yasir
;  APPLICANT:  Fling, Steve
;  APPLICANT:  Maisonneuve, Jeff
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE:  210121.469C5
;  CURRENT APPLICATION NUMBER: US/09/556,877
;  CURRENT FILING DATE:  2000-04-19
;  NUMBER OF SEQ ID NOS:  305
;  SOFTWARE:  FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 182
;   LENGTH: 3021
;   TYPE: DNA
;   ORGANISM: Chlamydia
US-09-556-877-182

  Query Match             1.7%;  Score 48.4;  DB 4;  Length 3021;
  Best Local Similarity   56.7%;  Pred. No. 0.001;
  Matches  110;  Conservative    0;  Mismatches   81;  Indels    3;  Gaps    1;

Qy     896 TCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTA 955
           | || ||||||  |||| |  |         | || ||||||||||| ||        |
Db    1073 TTTTCTTTAGTAGCAATGTAGCTGCTGGGAAAGGGGGAGCTATTTATGCCAAAAAGCTCT 1132

Qy     956 CCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCG 1015
           |  | |   |  |||||||    | ||| |||  |||||||| ||| ||   || || |
Db    1133 CGGTTGCTAACTGTGGCCCTGTACAATTTTTAAGGAATATCGCTAATGAT---GGTGGAG 1189

Qy    1016 CTATCTATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTT 1075
           | || ||| |||  | | |  |  |   ||   | ||||| ||        | ||||||||
Db    1190 CGATTTATTTAGGAGAATCTGGAGAGCTCAGTTTATCTGCTGATTATGGAGATATTATTT 1249

Qy    1076 TTAATGAAAATATT 1089
           |  |||  ||| ||
Db    1250 TCGATGGGAATCTT 1263


RESULT 8
US-09-620-412C-182
; Sequence 182, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 182
;   LENGTH: 3021
;   TYPE: DNA
;   ORGANISM: Chlamydia
US-09-620-412C-182

  Query Match             1.7%;  Score 48.4;  DB 4;  Length 3021;
  Best Local Similarity   56.7%;  Pred. No. 0.001;
  Matches  110;  Conservative    0;  Mismatches   81;  Indels    3;  Gaps    1;

Qy     896 TCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTA 955
           | || ||||||  |||| |  |         | || ||||||||||| ||        |
Db    1073 TTTTCTTTAGTAGCAATGTAGCTGCTGGGAAAGGGGGAGCTATTTATGCCAAAAAGCTCT 1132

Qy     956 CCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCG 1015
           |  | |   |  |||||||    | ||| |||  |||||||| ||| ||   || || |
Db    1133 CGGTTGCTAACTGTGGCCCTGTACAATTTTTAAGGAATATCGCTAATGAT---GGTGGAG 1189

Qy    1016 CTATCTATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTT 1075
           | || ||| |||  | | |  |  |   ||   | ||||| ||        | ||||||||
Db    1190 CGATTTATTTAGGAGAATCTGGAGAGCTCAGTTTATCTGCTGATTATGGAGATATTATTT 1249

Qy    1076 TTAATGAAAATATT 1089
           |  |||  ||| ||
Db    1250 TCGATGGGAATCTT 1263


RESULT 9
US-09-620-412C-316
; Sequence 316, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 316
;   LENGTH: 1941
;   TYPE: DNA
;   ORGANISM: Chlamydia trachomatis
US-09-620-412C-316

  Query Match             1.6%;  Score 45.2;  DB 4;  Length 1941;
  Best Local Similarity   49.2%;  Pred. No. 0.0066;
  Matches  239;  Conservative    0;  Mismatches  233;  Indels   14;  Gaps    4;

Qy     622 CAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCTGGAACGAGCAATTCTTTTGAG 681
```

```
Db   1114 CAAGCTGAAAATATGGCACGAGGAGGAGCAATCTGTAGTAGAAGAGACTTGTGCTCAATC 1173

Qy    682 AGTAATAACTGCGATC---TCTTCTTTATCAATAACGCCTGTTGTGCAGGAGGAGCGATC 738
Db   1174 AGCAATAATTCTGGTCCCATAGTTTTTAACTATAACCAAGGCGGGAAAGGTGGAGCTATT 1233

Qy    739 TTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATCGTTTTCTATAACAATCGC 798
Db   1234 AGCGCTACCCGATGTGTTATTGACAATAACAAAGAAAGAATCATCTTTTCAAACAATAGT 1293

Qy    799 TGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGATGGAGGAGCAATTAAAGTA 858
Db   1294 TCCCTGGGATGGAGCCAATCTTCTTCT------GCAAGTAACGGAGGAGCCATTCAAACG 1347

Qy    859 ACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTTTTTAGTGACAATATCACA 918
Db   1348 ACACAAGGATTTACTTTACGAAATAATAAAGGCTCTATCTACTTCGACAGCAACACTGCT 1407

Qy    919 AAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTAGTGGATAATGGCCCTACC 978
Db   1408 ACACACGCCGGGGGAGCCATTAACTGTGGTTACATTGACATCCGAGATAACGGACCCGTC 1467

Qy    979 TACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATCTATATAGACGGAACCAGC 1038
Db   1468 TATTTTCTAAATAA-CTCTGCTGCCTGGGGAGCGGCCTTTAATTTATCGAAACCACGTTC 1526

Qy   1039 AACTCCAAAATTTCTGCCGACCGCCAT----GCTATTATTTTTAATGAAAATATTGTGAC 1094
Db   1527 AGCGACAAATTATATCCATACAGGGACAGGCGATATTGTTTTTAATAATAACGTTGTCTT 1586

Qy   1095 TAATGT 1100
Db   1587 TACTCT 1592


RESULT 10
US-09-620-412C-324
; Sequence 324, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
;  APPLICANT: Steven P. Fling
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C7
;  CURRENT APPLICATION NUMBER: US/09/620,412C
;  CURRENT FILING DATE:  2000-07-20
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 324
;   LENGTH: 1896
;   TYPE: DNA
;   ORGANISM: Chlamydia trachomatis
US-09-620-412C-324

  Query Match             1.4%;  Score 41.4;  DB 4;  Length 1896;
  Best Local Similarity   54.2%;  Pred. No. 0.076;
  Matches   84;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy    869 TAGATGTTACAGGCAATCGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATG 928
Db   1370 TAAAGTTTGAAGGCAATAAAGGTTCTATTGTTTTTGATTACAACTTTGCAAAAGGCAGAG 1429

Qy    929 GCGGAGCTATTTACGCTCCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAA 988
Db   1430 GCGGAAGCATCCTAACGAAAGAATTCTCTCTTGTAGCAGATGATTCGGTTGTCTTTAGTA 1489

Qy    989 ACAATATCGCCAATAATAAGGGGGGCGCTATCTAT 1023
Db   1490 ACAATACAGCAGAAAAAGGCGGTGGAGCTATTTAT 1524


RESULT 11
```

```
US-08-915-107-3/c
; Sequence 3, Application US/08915107
; Patent No. 5885805
;  GENERAL INFORMATION:
;    APPLICANT:  Burnham, Martin K.R.
;    APPLICANT:  Lonetto, Michael A.
;    APPLICANT:  Warren, Patrick V.
;    TITLE OF INVENTION:  NOVEL GbpA
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dechert Price & Rhoads
;      STREET:  4000 Bell Atlantic Tower, 1717 Arch Stre
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  US
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/915,107
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dickinson, Todd Q
;      REGISTRATION NUMBER:  28,354
;      REFERENCE/DOCKET NUMBER:  P50549-4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-994-2252
;      TELEFAX:  215-994-2222
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1122 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-915-107-3

  Query Match             1.4%;  Score 40.6;  DB 2;  Length 1122;
  Best Local Similarity   46.3%;  Pred. No. 0.1;
  Matches  133;  Conservative    0;  Mismatches  154;  Indels    0;  Gaps    0;

Qy   1460 ATGGAGCAGTTCTGAGTTGCTATAAAAATGGTGCAGGAAATTCTGCTAGCAATGCCTCTA 1519
Db    943 ATTGAGGTGCAGTCATACCTTGTTTAAATGTCCAAGCACGTACTTCTTGCACACCAGCAG 884

Qy   1520 TAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTC 1579
Db    883 TAAAATATGTTGATAATCCTAATAATTCATAAGTTGTTCTAATTAATCGATCTAATCCTG 824

Qy   1580 CTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTATACAGCAGATACTGCAGCTA 1639
Db    823 GTTCTTCGATACCTAAATCTTCTAAGAACATTTCTTTATCTTCATCATCTAATGTAGCAA 764

Qy   1640 CCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGACTATGGGAATTCTCCTTATG 1699
Db    763 TTTCTTCTTCAATTTTTGCACTAATAACAATCACTTCAGAGTCTTCTTGCGCTGCATATT 704

Qy   1700 AATCCACAGATCTAACCCATGCTCTGTCATCACAGCCTATGCTATCT 1746
Db    703 CACGAATCGCTTTTACTTTATCATTATCATCATCACCAATTTCATCT 657


RESULT 12
US-09-273-613-3/c
; Sequence 3, Application US/09273613
; Patent No. 6203800
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  Burnham, Martin K.R.
;    APPLICANT:  Lonetto, Michael A.
;    APPLICANT:  Warren, Patrick V.
;    TITLE OF INVENTION:  NOVEL GbpA
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dechert Price & Rhoads
;      STREET:  4000 Bell Atlantic Tower, 1717 Arch Stre
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  US
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/273,613
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/915,107
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dickinson, Todd Q
;      REGISTRATION NUMBER:  28,354
;      REFERENCE/DOCKET NUMBER:  P50549-4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-994-2252
;      TELEFAX:  215-994-2222
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1122 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-09-273-613-3

  Query Match             1.4%;  Score 40.6;  DB 4;  Length 1122;
  Best Local Similarity   46.3%;  Pred. No. 0.1;
  Matches  133;  Conservative    0;  Mismatches  154;  Indels    0;  Gaps    0;

Qy   1460 ATGGAGCAGTTCTGAGTTGCTATAAAAATGGTGCAGGAAATTCTGCTAGCAATGCCTCTA 1519
          || |||  |   | |     | |  |||||    || |  | || || |||   |  |.
Db    943 ATTGAGGTGCAGTCATACCTTGTTTAAATGTCCAAGCACGTACTTCTTGCACACCAGCAG 884

Qy   1520 TAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTC 1579
          ||| |       ||| |      |||   |      |||| |  | | |  || |   |
Db    883 TAAAATATGTTGATAATCCTAATAATTCATAAGTTGTTCTAATTAATCGATCTAATCCTG 824

Qy   1580 CTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTATACAGCAGATACTGCAGCTA 1639
           ||  | |       |  || | || ||||    |   | | || ||  || || ||| |
Db    823 GTTCTTCGATACCTAAATCTTCTAAGAACATTTCTTTATCTTCATCATCTAATGTAGCAA 764

Qy   1640 CCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGACTATGGGAATTCTCCTTATG 1699
            | ||| | || |  || |  | |  || ||| |   || | |       || | |||
Db    763 TTTCTTCTTCAATTTTTGCACTAATAACAATCACTTCAGAGTCTTCTTGCGCTGCATATT 704

Qy   1700 AATCCACAGATCTAACCCATGCTCTGTCATCACAGCCTATGCTATCT 1746
          |   |  | | | ||     |  | ||||||   || ||   ||||
Db    703 CACGAATCGCTTTTACTTTATCATTATCATCATCACCAATTTCATCT 657


RESULT 13
US-09-310-867-3/c
; Sequence 3, Application US/09310867
; Patent No. 6320036
;  GENERAL INFORMATION:
;       APPLICANT: Burnham, Martin K.R.
;                  Lonetto, Michael A.
;                  Warren, Patrick V.
;       TITLE OF INVENTION: NOVEL GbpA
;       NUMBER OF SEQUENCES: 7
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Dechert Price & Rhoads
;            STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
;            CITY: Philadelphia
;            STATE: PA
;            COUNTRY: US
;            ZIP: 19103
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: FastSEQ for Windows Version 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/310,867
;            FILING DATE: 13-May-1999
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/915,107
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Dickinson, Todd Q
;            REGISTRATION NUMBER: 28,354
;            REFERENCE/DOCKET NUMBER: P50549-4
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 215-994-2252
;            TELEFAX: 215-994-2222
;            TELEX: <Unknown>
;  INFORMATION FOR SEQ ID NO: 3:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1122 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-310-867-3

  Query Match             1.4%;  Score 40.6;  DB 4;  Length 1122;
  Best Local Similarity   46.3%;  Pred. No. 0.1;
  Matches  133;  Conservative    0;  Mismatches  154;  Indels    0;  Gaps    0;

Qy   1460 ATGGAGCAGTTCTGAGTTGCTATAAAAATGGTGCAGGAAATTCTGCTAGCAATGCCTCTA 1519
          || |||  |   | |     | |  |||||    || |  | || || |||   |  |
Db    943 ATTGAGGTGCAGTCATACCTTGTTTAAATGTCCAAGCACGTACTTCTTGCACACCAGCAG 884

Qy   1520 TAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTC 1579
          ||| |       ||| |      |||   |      |||| |  | | |  || |   |
Db    883 TAAAATATGTTGATAATCCTAATAATTCATAAGTTGTTCTAATTAATCGATCTAATCCTG 824

Qy   1580 CTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTATACAGCAGATACTGCAGCTA 1639
           ||  | |       |  || | || ||||    |   | | || ||  || || ||| |
Db    823 GTTCTTCGATACCTAAATCTTCTAAGAACATTTCTTTATCTTCATCATCTAATGTAGCAA 764

Qy   1640 CCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGACTATGGGAATTCTCCTTATG 1699
            | ||| | || |  || |  | |  || ||| |   || | |       || | |||
Db    763 TTTCTTCTTCAATTTTTGCACTAATAACAATCACTTCAGAGTCTTCTTGCGCTGCATATT 704

Qy   1700 AATCCACAGATCTAACCCATGCTCTGTCATCACAGCCTATGCTATCT 1746
          |   |  | | | ||     |  | ||||||   || ||   ||||
Db    703 CACGAATCGCTTTTACTTTATCATTATCATCATCACCAATTTCATCT 657


RESULT 14
US-08-915-107-1/c
; Sequence 1, Application US/08915107
; Patent No. 5885805
;  GENERAL INFORMATION:
;    APPLICANT:  Burnham, Martin K.R.
;    APPLICANT:  Lonetto, Michael A.
;    APPLICANT:  Warren, Patrick V.
;    TITLE OF INVENTION:  NOVEL GbpA
```

```
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dechert Price & Rhoads
;      STREET:  4000 Bell Atlantic Tower, 1717 Arch Stre
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  US
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/915,107
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dickinson, Todd Q
;      REGISTRATION NUMBER:  28,354
;      REFERENCE/DOCKET NUMBER:  P50549-4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-994-2252
;      TELEFAX:  215-994-2222
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1125 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-915-107-1

  Query Match             1.4%;  Score 40.6;  DB 2;  Length 1125;
  Best Local Similarity   46.3%;  Pred. No. 0.1;
  Matches  133;  Conservative    0;  Mismatches  154;  Indels    0;  Gaps    0;

Qy   1460 ATGGAGCAGTTCTGAGTTGCTATAAAAATGGTGCAGGAAATTCTGCTAGCAATGCCTCTA 1519
          || |||  |   | |     | |  |||||    || |  | || || |||   |  |
Db    943 ATTGAGGTGCAGTCATACCTTGTTTAAATGTCCAAGCACGTACTTCTTGCACACCAGCAG 884

Qy   1520 TAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTC 1579
          |||  |       ||| |       |||   |       |||| |  | | |  || |   |
Db    883 TAAAATATGTTGATAATCCTAATAATTCATAAGTTGTTCTAATTAATCGATCTAATCCTG 824

Qy   1580 CTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTATACAGCAGATACTGCAGCTA 1639
           ||  | |       |  || | || ||||    |   | | || ||  || || ||| |
Db    823 GTTCTTCGATACCTAAATCTTCTAAGAACATTTCTTTATCTTCATCATCTAATGTAGCAA 764

Qy   1640 CCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGACTATGGGAATTCTCCTTATG 1699
            | ||| | || |  || |  | |  || ||| |   || | |       || | |||
Db    763 TTTCTTCTTCAATTTTTGCACTAATAACAATCACTTCAGAGTCTTCTTGCGCTGCATATT 704

Qy   1700 AATCCACAGATCTAACCCATGCTCTGTCATCACAGCCTATGCTATCT 1746
           |   |  | | | ||     |  | ||||||   || ||   ||||
Db    703 CACGAATCGCTTTTACTTTATCATTATCATCATCACCAATTTCATCT 657


RESULT 15
US-09-273-613-1/c
; Sequence 1, Application US/09273613
; Patent No. 6203800
;  GENERAL INFORMATION:
;    APPLICANT:  Burnham, Martin K.R.
;    APPLICANT:  Lonetto, Michael A.
;    APPLICANT:  Warren, Patrick V.
;    TITLE OF INVENTION:  NOVEL GbpA
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dechert Price & Rhoads
;      STREET:  4000 Bell Atlantic Tower, 1717 Arch Stre
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  US
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/273,613
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/915,107
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dickinson, Todd Q
;      REGISTRATION NUMBER:  28,354
;      REFERENCE/DOCKET NUMBER:  P50549-4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-994-2252
;      TELEFAX:  215-994-2222
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1125 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-09-273-613-1

  Query Match             1.4%;  Score 40.6;  DB 4;  Length 1125;
  Best Local Similarity   46.3%;  Pred. No. 0.1;
  Matches  133;  Conservative    0;  Mismatches  154;  Indels    0;  Gaps    0;

Qy   1460 ATGGAGCAGTTCTGAGTTGCTATAAAAATGGTGCAGGAAATTCTGCTAGCAATGCCTCTA 1519
          || |||  |   | |     | |  |||||    || |  | || || |||   |  |
Db    943 ATTGAGGTGCAGTCATACCTTGTTTAAATGTCCAAGCACGTACTTCTTGCACACCAGCAG 884

Qy   1520 TAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTC 1579
          |||  |       ||| |       |||   |       |||| |  | | |  || |   |
Db    883 TAAAATATGTTGATAATCCTAATAATTCATAAGTTGTTCTAATTAATCGATCTAATCCTG 824

Qy   1580 CTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTATACAGCAGATACTGCAGCTA 1639
           ||  | |       |  || | || ||||    |   | | || ||  || || ||| |
Db    823 GTTCTTCGATACCTAAATCTTCTAAGAACATTTCTTTATCTTCATCATCTAATGTAGCAA 764

Qy   1640 CCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGACTATGGGAATTCTCCTTATG 1699
            | ||| | || |  || |  | |  || ||| |   || | |       || | |||
Db    763 TTTCTTCTTCAATTTTTGCACTAATAACAATCACTTCAGAGTCTTCTTGCGCTGCATATT 704

Qy   1700 AATCCACAGATCTAACCCATGCTCTGTCATCACAGCCTATGCTATCT 1746
           |   |  | | | ||     |  | ||||||   || ||   ||||
Db    703 CACGAATCGCTTTTACTTTATCATTATCATCATCACCAATTTCATCT 657


Search completed: February  4, 2003, 19:10:13
Job time : 133 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 4, 2003, 18:21:32 ; Search time 172 Seconds
(without alignments)
7986.399 Million cell updates/sec

Title: US-09-677-752-1
Perfect score: 2898
Sequence: 1 atgaaaaaagcgtttttctt.....aaattgctctgcgattctag 2898

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 408267 seqs, 237001491 residues

Total number of hits satisfying chosen parameters: 816534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2821.2	97.3	2895	10	US-09-841-132-171	Sequence 171, App
	2	2777.8	95.9	2934	10	US-09-841-132-183	Sequence 183, App
c	3	1190.6	41.1	1634	12	US-10-007-693-43	Sequence 43, Appl
	4	115.8	4.0	3092	10	US-09-886-468-9	Sequence 9, Appli
	5	48.4	1.7	2949	10	US-09-841-132-170	Sequence 170, App
	6	48.4	1.7	3021	10	US-09-841-132-182	Sequence 182, App
	7	45.2	1.6	1941	10	US-09-841-132-316	Sequence 316, App
	8	41.6	1.4	504	10	US-09-847-539A-16	Sequence 16, Appl
	9	41.6	1.4	654	10	US-09-847-539A-12	Sequence 12, Appl
	10	41.6	1.4	764	10	US-09-847-539A-27	Sequence 27, Appl
	11	41.4	1.4	916	12	US-10-007-693-133	Sequence 133, App
	12	41.4	1.4	1862	12	US-10-007-693-44	Sequence 44, Appl
	13	41.4	1.4	1896	10	US-09-841-132-324	Sequence 324, App
	14	41.4	1.4	3048	12	US-10-007-693-80	Sequence 80, Appl
	15	40.8	1.4	966	10	US-09-815-242-8052	Sequence 8052, Ap
	16	40.8	1.4	966	10	US-09-815-242-8678	Sequence 8678, Ap
	17	40.8	1.4	966	10	US-09-815-242-8966	Sequence 8966, Ap
c	18	40.6	1.4	1095	10	US-09-815-242-4538	Sequence 4538, Ap
c	19	40.6	1.4	1098	10	US-09-815-242-8316	Sequence 8316, Ap
c	20	40.6	1.4	1098	10	US-09-815-242-8927	Sequence 8927, Ap
	21	40.6	1.4	1860	10	US-09-841-132-308	Sequence 308, App
	22	40	1.4	469	10	US-09-847-539A-14	Sequence 14, Appl
	23	39.6	1.4	5172	10	US-09-841-132-374	Sequence 374, App
	24	39.6	1.4	5172	10	US-09-841-132-375	Sequence 375, App
	25	38.8	1.3	1557	10	US-09-841-132-332	Sequence 332, App
	26	38.8	1.3	5253	10	US-09-841-132-421	Sequence 421, App
	27	38.8	1.3	5253	10	US-09-841-132-556	Sequence 556, App
	28	38.8	1.3	5265	10	US-09-841-132-174	Sequence 174, App
	29	38.4	1.3	777	10	US-09-847-539A-13	Sequence 13, Appl
	30	38.4	1.3	807	10	US-09-949-972-1	Sequence 1, Appli
	31	38.4	1.3	853	10	US-09-847-539A-15	Sequence 15, Appl
	32	38.4	1.3	2816	10	US-09-886-468-8	Sequence 8, Appli
	33	38.4	1.3	4224	10	US-09-841-132-486	Sequence 486, App
	34	38.2	1.3	249	9	US-10-046-935-1662	Sequence 1662, Ap
	35	38.2	1.3	249	9	US-09-878-178-1662	Sequence 1662, Ap
c	36	38.2	1.3	1377	10	US-09-764-864-152	Sequence 152, App
	37	38.2	1.3	4970	10	US-09-816-095-1	Sequence 1, Appli
	38	38.2	1.3	99916	10	US-09-816-095-3	Sequence 3, Appli
	39	38	1.3	2397	12	US-10-007-693-3	Sequence 3, Appli
	40	38	1.3	2547	10	US-09-841-132-184	Sequence 184, App
	41	38	1.3	4593	10	US-09-841-132-172	Sequence 172, App
	42	38	1.3	4593	12	US-10-007-693-83	Sequence 83, Appl
c	43	36.6	1.3	912	9	US-09-828-523A-19	Sequence 19, Appl
c	44	36.6	1.3	936	9	US-09-828-523A-85	Sequence 85, Appl
	45	36.2	1.2	640681	10	US-09-790-988-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-841-132-171
; Sequence 171, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 171
; LENGTH: 2895
; TYPE: DNA
; ORGANISM: Chlamydia
US-09-841-132-171

Query Match 97.3%; Score 2821.2; DB 10; Length 2895;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 2857; Conservative 0; Mismatches 38; Indels 3; Gaps 1;

```
Qy       1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATTGGAAACTCCCTATCAGGACTAGCTAGAGAG 60
           |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db       1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATCGGAAACTCCCTATCAGGACTAGCTAGAGAG 60

Qy      61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120

Qy     121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180

Qy     181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240

Qy     241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300
```

```
Db    241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300
Qy    301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360
Db    301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360
Qy    361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACCATTTACATCG 420
Db    361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACTATTTACATGG 420
Qy    421 AGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 480
Db    421 AGAAATCCTTATGCTGCTG---ATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 477
Qy    481 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCGA 540
Db    478 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCAA 537
Qy    541 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 600
Db    538 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 597
Qy    601 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 660
Db    598 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 657
Qy    661 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTTATCAATAACGCCTGT 720
Db    658 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTCATCAATAACGCCTGT 717
Qy    721 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 780
Db    718 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 777
Qy    781 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 840
Db    778 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 837
Qy    841 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 900
Db    838 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 897
Qy    901 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 960
Db    898 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 957
Qy    961 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1020
Db    958 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1017
Qy   1021 TATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1080
Db   1018 TATATAGACGGAACCAGTAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1077
Qy   1081 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1140
Db   1078 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1137
Qy   1141 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1200
Db   1138 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1197
Qy   1201 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 1260
Db   1198 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 1257
Qy   1261 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 1320
Db   1258 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 1317
Qy   1321 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 1380
Db   1318 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 1377
Qy   1381 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 1440
Db   1378 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 1437
Qy   1441 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTGCAGGAAAT 1500
Db   1438 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTACAGGAGAT 1497
Qy   1501 TCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTG 1560
Db   1498 TCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTG 1557
Qy   1561 AAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTAT 1620
Db   1558 AAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTAT 1617
Qy   1621 ACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGAC 1680
Db   1618 ACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGAC 1677
Qy   1681 TATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTGTCATCACAGCCTATG 1740
Db   1678 TACGGGAACTCTCCTTATGAATCCACAGATCTGACCCATGCTCTGTCATCACAGCCTATG 1737
Qy   1741 CTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGATATGGATTTTTCGGGA 1800
Db   1738 CTATCTATTTCTGAAGCTAGCGATAACCAGCTACAATCAGAAAATATAGATTTTTCGGGA 1797
Qy   1801 CTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAA 1860
Db   1798 CTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAA 1857
Qy   1861 GATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAAGCCAATAGATTCCAT 1920
Db   1858 GATCCAGAACCAGCATCTTCAGCAACAATCACTGATCCACAAAAAGCCAATAGATTTCAT 1917
Qy   1921 AGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGCCCGAAACACAGAAGT 1980
Db   1918 AGAACCTTACTACTAACATGGCTTCCTGCCGGGTATGTTCCTAGCCCAAAACACAGAAGT 1977
Qy   1981 CCCCTCATAGCGAATACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAAT 2040
Db   1978 CCCCTCATAGCTAACACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAAT 2037
Qy   2041 AGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 2100
Db   2038 AGTGCAGAGCTGACACCTAGTGGTCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 2097
Qy   2101 ATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGA 2160
Db   2098 ATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGA 2157
Qy   2161 TACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTGAAATTCAGTCAGACC 2220
Db   2158 TACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTGAAATTCAGTCAGACC 2217
Qy   2221 TACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGC 2280
Db   2218 TACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGC 2277
Qy   2281 CAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTT 2340
Db   2278 CAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTT 2337
Qy   2341 TACAGCTATGGAGACCATAACTGTCACCATTTCTATACCCAAGGAGAAAATCTAACATCT 2400
Db   2338 TACAGCTATGGAGACCATAACTGTCACCATTTCTATACTCAAGGAGAAAATCTAACATCT 2397
Qy   2401 CAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTGTTTTTTTTGATCTCCCTATGAAA 2460
Db   2398 CAAGGGACGTTCCGCAGTCAAACGATGGGAGGTGCTGTCTTTTTTGATCTCCCTATGAAA 2457
```

```
Qy   2461 CCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2458 CCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCT 2517

Qy   2521 AGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCT 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2518 AGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCT 2577

Qy   2581 TTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCAAAGA 2640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   2578 TTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCACAGA 2637

Qy   2641 CCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACCAGGG 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2638 CCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACCAGGG 2697

Qy   2701 ATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCG 2760
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   2698 ATCGCGACCCAGCTCCTAGCCAGTAAAGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCG 2757

Qy   2761 CGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTC 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2758 CGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTC 2817

Qy   2821 CATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAA 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2818 CATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAA 2877

Qy   2881 ATTGCTCTGCGATTCTAG 2898
          ||||||||||||||||||
Db   2878 ATTGCTCTGCGATTCTAG 2895
```

RESULT 2
US-09-841-132-183
; Sequence 183, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 183
; LENGTH: 2934
; TYPE: DNA
; ORGANISM: Chlamydia
US-09-841-132-183

Query Match 95.9%; Score 2777.8; DB 10; Length 2934;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 2813; Conservative 0; Mismatches 37; Indels 3; Gaps 1;

```
Qy     46 GGACTAGCTAGAGAGGTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCT 105
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     85 GGACTAGCTAGAGAGGTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCT 144

Qy    106 ACGAAAGAGTCGCTATCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAAC 165
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    145 ACGAAAGAGTCGCTATCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAAC 204

Qy    166 TGCTATCTCGATAACCTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGA 225
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    205 TGCTATCTCGATAACCTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGA 264

Qy    226 GCTGCTGTCACAATAACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTAT 285
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    265 GCTGCTGTCACAATAACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTAT 324

Qy    286 TTTGCAAAAAATCTCACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCT 345
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    325 TTTGCAAAAAATCTCACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCT 384

Qy    346 CCTACCGTGGAGATTCGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGC 405
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    385 CCTACCGTGGAGATTCGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGC 444

Qy    406 AGACCATTTACATCGAGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCC 465
          |||| |||||||| ||| |||||| |||| ||||   ||||||||||||||||||||||
Db    445 AGACTATTTACATGGAGAAATCCTTATGCTGCTG---ATAAAATAAGAGAAGGCGGAGCC 501

Qy    466 ATTCATGCTCAAAATCTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAAC 525
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    502 ATTCATGCTCAAAATCTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAAC 561

Qy    526 TTTTCTTATGTCCGAGGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAAT 585
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    562 TTTTCTTATGTCCAAGGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAAT 621

Qy    586 CAGTCTTGTTTTCTCTTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGT 645
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    622 CAGTCTTGTTTTCTCTTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGT 681

Qy    646 GGCGCTATCTATGCTGGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTT 705
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    682 GGCGCTATCTATGCTGGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTC 741

Qy    706 ATCAATAACGCCTGTTGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGA 765
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    742 ATCAATAACGCCTGTTGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGA 801

Qy    766 AATCGTGGTAACATCGTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCT 825
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    802 AATCGTGGTAACATCGTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCT 861

Qy    826 TCAGAAGCTTCTGATGGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAAT 885
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    862 TCAGAAGCTTCTGATGGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAAT 921

Qy    886 CGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCT 945
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    922 CGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCT 981

Qy    946 CCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAAT 1005
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    982 CCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAAT 1041

Qy   1006 AAGGGGGGCGCTATCTATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCAT 1065
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   1042 AAGGGGGGCGCTATCTATATAGACGGAACCAGTAACTCCAAAATTTCTGCCGACCGCCAT 1101

Qy   1066 GCTATTATTTTTAATGAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACG 1125
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1102 GCTATTATTTTTAATGAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACG 1161

Qy   1126 TCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTA 1185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1162 TCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTA 1221

Qy   1186 TTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCA 1245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1222 TTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCA 1281

Qy   1246 GGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGA 1305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1282 GGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGA 1341

Qy   1306 GCTACTGTTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCC 1365
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1342 GCTACTGTTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCC 1401
```

```
Qy   1366 CTTACTCTCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGA 1425
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1402 CTTACTCTCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGA 1461

Qy   1426 TTCACACAAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAA 1485
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1462 TTCACACAAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAA 1521

Qy   1486 AATGGTGCAGGAAATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAAT 1545
          |||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   1522 AATGGTACAGGAGATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAAT 1581

Qy   1546 CTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAAT 1605
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1582 CTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAAT 1641

Qy   1606 AACAGCAATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTC 1665
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1642 AACAGCAATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTC 1701

Qy   1666 TCACTCATTGATGACTATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTG 1725
          |||||||||||||||| ||||| ||||||||||||||||||||| ||||||||||||
Db   1702 TCACTCATTGATGACTACGGGAACTCTCCTTATGAATCCACAGATCTGACCCATGCTCTG 1761

Qy   1726 TCATCACAGCCTATGCTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGAT 1785
          |||||||||||||||||||||||||| ||||| |||||||||||  ||| ||  ||
Db   1762 TCATCACAGCCTATGCTATCTATTTCTGAAGCTAGCGATAACCAGCTACAATCAGAAAAT 1821

Qy   1786 ATGGATTTTTCGGGACTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGC 1845
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1822 ATAGATTTTTCGGGACTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGC 1881

Qy   1846 TGGGCAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAA 1905
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db   1882 TGGGCAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACTGATCCACAAAAA 1941

Qy   1906 GCCAATAGATTCCATAGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGC 1965
          |||||||||||| |||||||||||| |||| || |||||||||||| ||||||||||||||
Db   1942 GCCAATAGATTTCATAGAACCTTACTACTAACATGGCTTCCTGCCGGGTATGTTCCTAGC 2001

Qy   1966 CCGAAACACAGAAGTCCCCTCATAGCGAATACCTTATGGGGGAATATGCTGCTTGCAACA 2025
          || ||||||||||||||||||||||| || ||||||||||||||||||||||||||||||
Db   2002 CCAAAACACAGAAGTCCCCTCATAGCTAACACCTTATGGGGGAATATGCTGCTTGCAACA 2061

Qy   2026 GAAAGCTTAAAAAATAGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTACA 2085
          |||||||||||||||||||||| |||||||||||| ||||||||||||||||||||||||
Db   2062 GAAAGCTTAAAAAATAGTGCAGAGCTGACACCTAGTGGTCATCCTTTCTGGGGAATTACA 2121

Qy   2086 GGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCAT 2145
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2122 GGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCAT 2181

Qy   2146 ATGCGCTCTTCCGGATACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTG 2205
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2182 ATGCGCTCTTCCGGATACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTG 2241

Qy   2206 AAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCT 2265
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2242 AAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCT 2301

Qy   2266 AAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACT 2325
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2302 AAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACT 2361

Qy   2326 AAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCACCATTTCTATACCCAAGGA 2385
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   2362 AAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCACCATTTCTATACTCAAGGA 2421

Qy   2386 GAAAATCTAACATCTCAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTGTTTTTTTT 2445
          |||||||||||||||||||||||| || |||||||||||||||||||| ||||||
Db   2422 GAAAATCTAACATCTCAAGGGACGTTCCGCAGTCAAACGATGGGAGGTGCTGTCTTTTTT 2481
```

```
Qy   2446 GATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCT 2505
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2482 GATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCT 2541

Qy   2506 CTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTT 2565
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2542 CTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTT 2601

Qy   2566 TCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATG 2625
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2602 TCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATG 2661

Qy   2626 AATGCTACCCAAAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTAT 2685
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   2662 AATGCTACCCACAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTAT 2721

Qy   2686 AGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGTTTGGTAGT 2745
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   2722 AGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAGTAAAGGTATTTGGTTTGGTAGT 2781

Qy   2746 GGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTG 2805
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2782 GGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTG 2841

Qy   2806 AGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAAT 2865
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2842 AGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAAT 2901

Qy   2866 TATCTCAATGGGGAAATTGCTCTGCGATTCTAG 2898
          ||||||||||||||||||||||||||||||||
Db   2902 TATCTCAATGGGGAAATTGCTCTGCGATTCTAG 2934
```

RESULT 3
US-10-007-693-43/c
; Sequence 43, Application US/10007693
; Patent No. US20020146776A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515C2
; CURRENT APPLICATION NUMBER: US/10/007,693
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 43
; LENGTH: 1634
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-10-007-693-43

Query Match 41.1%; Score 1190.6; DB 12; Length 1634;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1209; Conservative 0; Mismatches 9; Indels 3; Gaps 1;

```
Qy      1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATTGGAAACTCCCTATCAGGACTAGCTAGAGAG 60
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   1218 ATGAAAAAAGCGTTTTTCTTTTTCCTTATCGGAAACTCCCTATCAGGACTAGCTAGAGAG 1159

Qy     61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1158 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 1099

Qy    121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1098 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 1039

Qy    181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1038 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 979

Qy    241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300
```

```
Db    978 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 919

Qy    301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360
Db    918 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 859

Qy    361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACCATTTACATCG 420
Db    858 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACTATTTACATGG 799

Qy    421 AGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 480
Db    798 AGAAATCCTTATGCTGCTG---ATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 742

Qy    481 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCGA 540
Db    741 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCAA 682

Qy    541 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 600
Db    681 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 622

Qy    601 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 660
Db    621 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 562

Qy    661 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTTATCAATAACGCCTGT 720
Db    561 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTCATCAATAACGCCTGT 502

Qy    721 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 780
Db    501 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 442

Qy    781 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 840
Db    441 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 382

Qy    841 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 900
Db    381 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 322

Qy    901 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 960
Db    321 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 262

Qy    961 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1020
Db    261 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 202

Qy   1021 TATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1080
Db    201 TATATAGACGGAACCAGTAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 142

Qy   1081 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1140
Db     141 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 82

Qy   1141 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1200
Db      81 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 22

Qy   1201 AGCCAAAATTTAATTTTTTAT 1221
Db      21 AGCCAAAATTTAATTTTTTAT 1

RESULT 4
US-09-886-468-9
; Sequence 9, Application US/09886468
; Patent No. US20020037293A1
; GENERAL INFORMATION:
;  APPLICANT: Aventis Pasteur Limited
;  TITLE OF INVENTION: Chlamydia antigens and corresponding DNA fragments and uses t
;  FILE REFERENCE: 77813-5
;  CURRENT APPLICATION NUMBER: US/09/886,468
;  CURRENT FILING DATE:  1999-12-23
;  PRIOR APPLICATION NUMBER: 60/113,280
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: 60/113,281
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: 60/113,282
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: 60/113,283
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: 60/113,284
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: 60/113,285
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: 60/113,385
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: 60/114,050
;  PRIOR FILING DATE: 1998-12-28
;  PRIOR APPLICATION NUMBER: 60/114,056
;  PRIOR FILING DATE: 1998-12-28
;  PRIOR APPLICATION NUMBER: 60/114,057
;  PRIOR FILING DATE: 1998-12-28
;  PRIOR APPLICATION NUMBER: 60/114,058
;  PRIOR FILING DATE: 1998-12-28
;  PRIOR APPLICATION NUMBER: 60/114,059
;  PRIOR FILING DATE: 1998-12-28
;  PRIOR APPLICATION NUMBER: 60/114,061
;  PRIOR FILING DATE: 1998-12-28
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 3092
;   TYPE: DNA
;   ORGANISM: Chlamydia pneumoniae
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (101)..(2989)
US-09-886-468-9

  Query Match             4.0%;  Score 115.8;  DB 10;  Length 3092;
  Best Local Similarity   44.3%;  Pred. No. 1.3e-21;
  Matches  792;  Conservative    0;  Mismatches  947;  Indels   48;  Gaps    6;

Qy   1117 ACCAGTACGTCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGT 1176
Db   1247 ACATCTTCTTCTCCTCAACCCGGATATAGAAATGCACTCTATGCTGCTCCGGGGATTAAC 1306

Qy   1177 GAAATTCTATTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAAGTT 1236
Db   1307 TTAAAACTAGGAGCAAGACAGGGTTATAAAATTCTCTTTTATGATCCTATAGATCACGAT 1366

Qy   1237 AGCAATGCAGGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTA 1296
Db   1367 CAGACGACAACAGATCCTATAGTATTTAATTATGAACCCCATCACCTTGGCACCGTGTTG 1426

Qy   1297 TTTTCAGGAGCTACTGTTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACA 1356
Db   1427 TTTTCCGGAATCAATGTAGATTCTAACGCAACAAATCCATTGAACTTCCTATCAAAATTT 1486

Qy   1357 CCTGCACCCCTTACTCTCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACA 1416
Db   1487 TCTAACTCTTCACGACTTGAAAGGGGTGTGCTCGCTATTGAAGATCGGGCTGCTATTTCT 1546

Qy   1417 GTGAATCGATTCACACAAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGT 1476
Db   1547 TGCAAAACCCTATCGCAAACTGGGGGCATTCTACGTTTAGGAAACGCAGCATTAATCAG- 1605

Qy   1477 TGCTATAAAAATGGTGCAGGAAATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATT 1536
Db   1606 -----GACGAAAGGCCCGGGAAG--------------CTCCATAAATTTTAATGCAATC 1645
```

```
Qy   1537 GGATTGAATCTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAG 1596
Db   1646 GCGATCAATCTTCCTTCTATTTTACAATCAGAAGCCTCAGCTCCAAAGTTCTGGATTTAT 1705

Qy   1597 CCTACAAATAACAGCAATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGTGAT 1656
Db   1706 CCTACATTAACAGGATCCACCTATTCTGAAGACACTTCTTCTACTATCACTCTCTCAGGA 1765

Qy   1657 GTAAAACTCTCACTCATTGATGACTATGGGAATTCTCCTTATGAATCCACAGATCTAACC 1716
Db   1766 CCCTTGACTTTTCTAAACGATGAAAATGAAAACCCCTATGATAGCTTAGATCTCTCTGAA 1825

Qy   1717 CATGCTCTGTCATCACAGCCTATGCTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGA 1776
Db   1826 CCTCGAAAGGATATCCCCCCTCCTCTACCTCCTCGATGTGACTGCAAAAAAATCGATACT 1885

Qy   1777 TCTGATGATATGGATTTTTCGGGACTAAATGTCCCTCATTATGGATGGCAAGGACTTTGG 1836
Db   1886 TCGAATCTCATTGTAGAAGCCATGAACTTAGATGAGCACTATGGATATCAGGGAATCTGG 1945

Qy   1837 ACTTGGGGCTGGGCAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACAGAT 1896
Db   1946 TCTCCCTATTGGATGGAAACT--------------ACGACTACAACAAGCTCTACAGTA 1990

Qy   1897 CCACAAAAAGCCAATAGATTCCATAGAACCTTATTACTGACTTGGCTTCCTGCTGGGTAT 1956
Db   1991 CCGGAACAGACCAATACAAACCACAGGCAGCTCTACGTAGACTGGACTCCTGTAGGATAC 2050

Qy   1957 GTTCCTAGCCCGAAACACAGAAGTCCCCTCATAGCGAATACCTTATGGGGGAATATGCTG 2016
Db   2051 CGCCCTAACCCGGAACGTCACGGAGAATTTATTGCTAATACCTTATGGCAGTCTGCCTAT 2110

Qy   2017 CTTGCAACAGAAAGCTTAAAAAATAGTGCAGAACTGACACCTAGTGATCATCCTTTCTGG 2076
Db   2111 AACGCTCTGTTAGGAATCCGCATCTTACCTCCACAAAACCTCAAAGAGCATGACCT---T 2167

Qy   2077 GGAATTACAGGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCATCCT 2136
Db   2168 GAAGCCTCTCTGCAAGGACTCGGGCTTCTAATTAACCAACATAATCGCGAGGGACGCAAA 2227

Qy   2137 GGATTCCATATGCGCTCTTCCGGATACTCTGCGGGGATGATAGCAGGGCAGAC------A 2190
Db   2228 GGCTTCCGAAACCATACTACGGGCTATGCAGCAACAACCTCAGCAAAAACTGCAGCACGA 2287

Qy   2191 CACACCTTCTCATTGAAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTACGCAAAA 2250
Db   2288 CATAGTTTCTCTTTAGGATTCGCACAAATGTTCTCCAAAACTAGAGAACGTCAATCTCCA 2347

Qy   2251 AACAACGTATCTTCTAAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAA 2310
Db   2348 AGTACGACTTCCTCCCACAACTACTTTGCAGGACTCCGCTTCGACAGTCTCCTCTTCAGG 2407

Qy   2311 GGTTTCTTGCTGACTAAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCACCAT 2370
Db   2408 GACTTCATCTCTACAGGGCTATCCCTAGGTTATAGCTACGGAGATCACCATATGCTTTGC 2467

Qy   2371 TTCTATACCCAAGGAGAAAATCTAACATCTCAAGGGACGTTCCGTAGTCAAACGATGGGA 2430
Db   2468 CACTATACAGAAATCTTAAAAGGGTCGTCCAAAGCCTTCTTTAATAACCACACTTTGGTA 2527

Qy   2431 GGTGCTGTTTTTTTTGATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCT 2490
Db   2528 GCCTCTCTAGACTGCACATTCTTACCAGCTAGAATCACCCGCACTCTCGAACTC---CAG 2584

Qy   2491 CCCTTTTTAGGTGCTCTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTGGGAGCC 2550
Db   2585 CCCTTTATCAGTGCCATTGCTCTGCGCTGTTCCCAGGCCTCGTTCCAAGAAACTGGAGAC 2644

Qy   2551 TATCCGCGAAGCTTTTCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTT 2610
Db   2645 CATATAAGAAAATTCCATCCAAAACATCCCCTTACAGATCTTTCCTCTCCCATAGGCTTC 2704
```

```
Qy   2611 AAAGGTAGCTTTATGAATGCTACCCAAAGACCTCAAGCCTGGACTGTAGAATTGGCATAC 2670
Db   2705 CGTTCTGAATGGAAAACTTCACATCATATCCCCATGCTATGGACTACGGAAATATCCTAC 2764

Qy   2671 CAACCCGTTCTGTATAGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAGTAAGGGT 2730
Db   2765 GTACCTACCCTATACAGAAAAAATCCAGAAATGTTCACGACACTACTCATCAGCAATGGA 2824

Qy   2731 ATTTGGTTTGGTAGTGGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAG 2790
Db   2825 ACATGGACAACACAAGCAACTCCCGTCTCCTATAATTCCGTAGCTGCAAAAATAAAAAAT 2884

Qy   2791 CAAACACAACCTTTGAGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCT 2850
Db   2885 ACTTCCCAACTTTTCTCAAGAGTAACCTTATCCTTAGATTATTCAGCTCAAGTCTCCTCG 2944

Qy   2851 TCAACCTTCTGTAATTATCTCAATGGGGAAATTGCTCTGCGATTCTA 2897
Db   2945 TCAACTGTAGGTCAATACCTTAAAGCTGAGAGTCATTGCACATTTTA 2991
```

RESULT 5
US-09-841-132-170
; Sequence 170, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 170
; LENGTH: 2949
; TYPE: DNA
; ORGANISM: Chlamydia
US-09-841-132-170

Query Match 1.7%; Score 48.4; DB 10; Length 2949;
Best Local Similarity 56.7%; Pred. No. 0.0049;
Matches 110; Conservative 0; Mismatches 81; Indels 3; Gaps 1;

```
Qy    896 TCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTA 955
Db   1001 TTTTCTTTAGTAGCAATGTAGCTGCTGGGAAAGGGGGAGCTATTTATGCCAAAAAGCTCT 1060

Qy    956 CCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCG 1015
Db   1061 CGGTTGCTAACTGTGGCCCTGTACAATTTTTAAGGAATATCGCTAATGAT---GGTGGAG 1117

Qy   1016 CTATCTATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTT 1075
Db   1118 CGATTTATTTAGGAGAATCTGGAGAGCTCAGTTTATCTGCTGATTATGGAGATATTATTT 1177

Qy   1076 TTAATGAAAATATT 1089
Db   1178 TCGATGGGAATCTT 1191
```

RESULT 6
US-09-841-132-182
; Sequence 182, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION

```
;  FILE REFERENCE: 210121.469C8
;  CURRENT APPLICATION NUMBER: US/09/841,132
;  CURRENT FILING DATE:  2001-04-23
;  NUMBER OF SEQ ID NOS: 599
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 182
;   LENGTH: 3021
;   TYPE: DNA
;   ORGANISM: Chlamydia
US-09-841-132-182

  Query Match             1.7%;  Score 48.4;  DB 10;  Length 3021;
  Best Local Similarity   56.7%;  Pred. No. 0.005;
  Matches  110;  Conservative    0;  Mismatches   81;  Indels    3;  Gaps    1;

Qy     896 TCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTA 955
           | || ||||||  |||| |  |        | || ||||||||||| ||        |
Db    1073 TTTTCTTTAGTAGCAATGTAGCTGCTGGGAAAGGGGGAGCTATTTATGCCAAAAAGCTCT 1132

Qy     956 CCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCG 1015
           |  | |   |  |||||||   | ||| |||  |||||||| ||| ||  || || |
Db    1133 CGGTTGCTAACTGTGGCCCTGTACAATTTTTAAGGAATATCGCTAATGAT---GGTGGAG 1189

Qy    1016 CTATCTATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTT 1075
           | || ||| |||  | | |  |  |   ||   | ||||| ||       | ||||||||
Db    1190 CGATTTATTTAGGAGAATCTGGAGAGCTCAGTTTATCTGCTGATTATGGAGATATTATTT 1249

Qy    1076 TTAATGAAAATATT 1089
           |  |||  ||| ||
Db    1250 TCGATGGGAATCTT 1263


RESULT 7
US-09-841-132-316
; Sequence 316, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
;  APPLICANT: Bhatia, Ajay
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C8
;  CURRENT APPLICATION NUMBER: US/09/841,132
;  CURRENT FILING DATE:  2001-04-23
;  NUMBER OF SEQ ID NOS: 599
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 316
;   LENGTH: 1941
;   TYPE: DNA
;   ORGANISM: Chlamydia trachomatis
US-09-841-132-316

  Query Match             1.6%;  Score 45.2;  DB 10;  Length 1941;
  Best Local Similarity   49.2%;  Pred. No. 0.03;
  Matches  239;  Conservative    0;  Mismatches  233;  Indels   14;  Gaps    4;

Qy     622 CAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCTGGAACGAGCAATTCTTTTGAG 681
           ||| || | |     | |  ||| || || |||| |  | |||    |   |  |
Db    1114 CAAGCTGAAAATATGGCACGAGGAGGAGCAATCTGTAGTAGAAGAGACTTGTGCTCAATC 1173

Qy     682 AGTAATAACTGCGATC---TCTTCTTTATCAATAACGCCTGTTGTGCAGGAGGAGCGATC 738
           || ||||| |  | ||    |  | |||| | |||||    |  |   ||| ||||| ||
Db    1174 AGCAATAATTCTGGTCCCATAGTTTTTAACTATAACCAAGGCGGGAAAGGTGGAGCTATT 1233

Qy     739 TTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATCGTTTTCTATAACAATCGC 798
             | |  |    |||  | |       ||    |  |  ||| | || |  |||||| |
Db    1234 AGCGCTACCCGATGTGTTATTGACAATAACAAAGAAAGAATCATCTTTTCAAACAATAGT 1293

Qy     799 TGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGATGGAGGAGCAATTAAAGTA 858
           | | |   |       || |  |||||      ||   | | |||||||| ||| ||
Db    1294 TCCCTGGGATGGAGCCAATCTTCTTCT------GCAAGTAACGGAGGAGCCATTCAAACG 1347

Qy     859 ACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTTTTTAGTGACAATATCACA 918
           ||      |  |   | |   |   |||   ||    ||||  ||       ||| |   |
Db    1348 ACACAAGGATTTACTTTACGAAATAATAAAGGCTCTATCTACTTCGACAGCAACACTGCT 1407

Qy     919 AAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTAGTGGATAATGGCCCTACC 978
           | | |    || ||||| ||| ||  |  |    ||  | |    |||||  || ||   |
Db    1408 ACACACGCCGGGGGAGCCATTAACTGTGGTTACATTGACATCCGAGATAACGGACCCGTC 1467

Qy     979 TACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATCTATATAGACGGAACCAGC 1038
           || ||| |||| ||  ||  |         | | |||  |   | ||| ||    ||   |
Db    1468 TATTTTCTAAATAA-CTCTGCTGCCTGGGGAGCGGCCTTTAATTTATCGAAACCACGTTC 1526

Qy    1039 AACTCCAAAATTTCTGCCGACCGCCAT----GCTATTATTTTTAATGAAAATATTGTGAC 1094
           | |  |||| | | | |  || | |      | |||| |||||||| | ||  ||||
Db    1527 AGCGACAAATTATATCCATACAGGGACAGGCGATATTGTTTTTAATAATAACGTTGTCTT 1586

Qy    1095 TAATGT 1100
           || | |
Db    1587 TACTCT 1592


RESULT 8
US-09-847-539A-16
; Sequence 16, Application US/09847539A
; Patent No. US20020061306A1
; GENERAL INFORMATION:
;  APPLICANT: Bjorck, Lars H
;  APPLICANT:  Rassmussen, Magnus
;  TITLE OF INVENTION: STREPTOCOCCAL ALPHA ZM BINDING PROTEIN
;  FILE REFERENCE: 100084.415US / N.75312B
;  CURRENT APPLICATION NUMBER: US/09/847,539A
;  CURRENT FILING DATE:  2001-05-01
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
;   LENGTH: 504
;   TYPE: DNA
;   ORGANISM: Streptococcus pyogenes
US-09-847-539A-16

  Query Match             1.4%;  Score 41.6;  DB 10;  Length 504;
  Best Local Similarity   50.0%;  Pred. No. 0.14;
  Matches  104;  Conservative    0;  Mismatches  104;  Indels    0;  Gaps    0;

Qy    1119 CAGTACGTCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGA 1178
           | | || | | |||||| || | | |    | ||  || | | |||        ||  ||
Db      75 CGGAACCTTAACTAATCTTCTTGGCAATGCTCCAGAAAAACTGGCATTACGTAATGAAGA 134

Qy    1179 AATTCTATTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAG 1238
           || *    | |  | |    |  | |||  | |       | | ||  ||| |  || ||
Db     135 AAGAGCCATTGATGAATTAAAAAAACAAGCTATTGAGGATAAAGAAGCTACGACAGCTAT 194

Qy    1239 CAATGCAGGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATT 1298
             | ||||     || |   ||||  |       ||| ||||||||||  | || ||  ||
Db     195 AGAAGCAGCAAGTTCAGATGCCTTAGAAGCATTAGCGGATCAAACAGACGCTTTACAATC 254

Qy    1299 TTCAGGAGCTACTGTTAATTCTGCAGAT 1326
              || |||| | |||  |   || |||
Db     255 AGAAGAAGCTGCGGTTGTTAAAGCGGAT 282


RESULT 9
US-09-847-539A-12
; Sequence 12, Application US/09847539A
; Patent No. US20020061306A1
; GENERAL INFORMATION:
;  APPLICANT: Bjorck, Lars H
;  APPLICANT:  Rassmussen, Magnus
;  TITLE OF INVENTION: STREPTOCOCCAL ALPHA ZM BINDING PROTEIN
;  FILE REFERENCE: 100084.415US / N.75312B
;  CURRENT APPLICATION NUMBER: US/09/847,539A
;  CURRENT FILING DATE:  2001-05-01
;  NUMBER OF SEQ ID NOS: 30
```

```
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
;   LENGTH: 654
;   TYPE: DNA
;   ORGANISM: Streptococcus pyogenes
US-09-847-539A-12

  Query Match             1.4%;  Score 41.6;  DB 10;  Length 654;
  Best Local Similarity   50.0%;  Pred. No. 0.16;
  Matches  104;  Conservative    0;  Mismatches  104;  Indels    0;  Gaps    0;

Qy   1119 CAGTACGTCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGA 1178
          | | || | | |||||| || | | |    | ||  || | | |||        ||  ||
Db    141 CGGAACCTTAACTAATCTTCTTGGCAATGCTCCAGAAAAACTGGCATTACGTAATGAAGA 200

Qy   1179 AATTCTATTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAG 1238
          ||      | |  | |   |  | ||| | |      | | ||  |||    || ||
Db    201 AAGAGCCATTGATGAATTAAAAAAACAAGCTATTGAGGATAAAGAAGCTACGACAGCTAT 260

Qy   1239 CAATGCAGGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATT 1298
            | ||||     || |   |||| |      ||| ||||||||||| | || ||  ||
Db    261 AGAAGCAGCAAGTTCAGATGCCTTAGAAGCATTAGCGGATCAAACAGACGCTTTACAATC 320

Qy   1299 TTCAGGAGCTACTGTTAATTCTGCAGAT 1326
             || |||| | |||  |   || |||
Db    321 AGAAGAAGCTGCGGTTGTTAAAGCGGAT 348


RESULT 10
US-09-847-539A-27
; Sequence 27, Application US/09847539A
; Patent No. US20020061306A1
; GENERAL INFORMATION:
;  APPLICANT: Bjorck, Lars H
;  APPLICANT:  Rassmussen, Magnus
;  TITLE OF INVENTION: STREPTOCOCCAL ALPHA ZM BINDING PROTEIN
;  FILE REFERENCE: 100084.415US / N.75312B
;  CURRENT APPLICATION NUMBER: US/09/847,539A
;  CURRENT FILING DATE:  2001-05-01
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
;   LENGTH: 764
;   TYPE: DNA
;   ORGANISM: Streptococcus pyogenes
US-09-847-539A-27

  Query Match             1.4%;  Score 41.6;  DB 10;  Length 764;
  Best Local Similarity   50.0%;  Pred. No. 0.17;
  Matches  104;  Conservative    0;  Mismatches  104;  Indels    0;  Gaps    0;

Qy   1119 CAGTACGTCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGA 1178
          | | || | | |||||| || | | |    | ||  || | | |||        ||  ||
Db    200 CGGAACCTTAACTAATCTTCTTGGCAATGCTCCAGAAAAACTGGCATTACGTAATGAAGA 259

Qy   1179 AATTCTATTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAG 1238
          ||      | |  | |   |  | ||| | |      | | ||  |||    || ||
Db    260 AAGAGCCATTGATGAATTAAAAAAACAAGCTATTGAGGATAAAGAAGCTACGACAGCTAT 319

Qy   1239 CAATGCAGGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATT 1298
            | ||||     || |   |||| |      ||| ||||||||||| | || ||  ||
Db    320 AGAAGCAGCAAGTTCAGATGCCTTAGAAGCATTAGCGGATCAAACAGACGCTTTACAATC 379

Qy   1299 TTCAGGAGCTACTGTTAATTCTGCAGAT 1326
             || |||| | |||  |   || |||
Db    380 AGAAGAAGCTGCGGTTGTTAAAGCGGAT 407


RESULT 11
US-10-007-693-133
; Sequence 133, Application US/10007693
; Patent No. US20020146776A1
; GENERAL INFORMATION:
;  APPLICANT: Bhatia, Ajay
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
;  TITLE OF INVENTION:  AND DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.515C2
;  CURRENT APPLICATION NUMBER: US/10/007,693
;  CURRENT FILING DATE:  2001-12-05
;  NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 133
;   LENGTH: 916
;   TYPE: DNA
;   ORGANISM: Chlamydia trachomatis
US-10-007-693-133

  Query Match             1.4%;  Score 41.4;  DB 12;  Length 916;
  Best Local Similarity   54.2%;  Pred. No. 0.22;
  Matches   84;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy    869 TAGATGTTACAGGCAATCGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATG 928
          || |* ||  |||||||   |||   ||  |||||  | ||||  |  |||||      |
Db    730 TAAAGTTTGAAGGCAATAAAGGTTCTATTGTTTTTGATTACAACTTTGCAAAAGGCAGAG 789

Qy    929 GCGGAGCTATTTACGCTCCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAA 988
          |||||   ||     |    | | |  | || || |   |||   |      |||||  |
Db    790 GCGGAAGCATCCTAACGAAAGAATTCTCTCTTGTAGCAGATGATTCGGTTGTCTTTAGTA 849

Qy    989 ACAATATCGCCAATAATAAGGGGGGCGCTATCTAT 1023
          ||||||  ||  | ||    || || |||||  |||
Db    850 ACAATACAGCAGAAAAAGGCGGTGGAGCTATTTAT 884


RESULT 12
US-10-007-693-44
; Sequence 44, Application US/10007693
; Patent No. US20020146776A1
; GENERAL INFORMATION:
;  APPLICANT: Bhatia, Ajay
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
;  TITLE OF INVENTION:  AND DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.515C2
;  CURRENT APPLICATION NUMBER: US/10/007,693
;  CURRENT FILING DATE:  2001-12-05
;  NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 44
;   LENGTH: 1862
;   TYPE: DNA
;   ORGANISM: Chlamydia trachomatis
US-10-007-693-44

  Query Match             1.4%;  Score 41.4;  DB 12;  Length 1862;
  Best Local Similarity   54.2%;  Pred. No. 0.32;
  Matches   84;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy    869 TAGATGTTACAGGCAATCGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATG 928
          || |  ||  |||||||   |||   ||  |||||  | ||||  |  |||||      |
Db   1178 TAAAGTTTGAAGGCAATAAAGGTTCTATTGTTTTTGATTACAACTTTGCAAAAGGCAGAG 1237

Qy    929 GCGGAGCTATTTACGCTCCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAA 988
          |||||   ||     |    | | |  | || || |   |||   |      |||||  |
Db   1238 GCGGAAGCATCCTAACGAAAGAATTCTCTCTTGTAGCAGATGATTCGGTTGTCTTTAGTA 1297

Qy    989 ACAATATCGCCAATAATAAGGGGGGCGCTATCTAT 1023
          ||||||  ||  | ||    || || |||||  |||
Db   1298 ACAATACAGCAGAAAAAGGCGGTGGAGCTATTTAT 1332


RESULT 13
US-09-841-132-324
; Sequence 324, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
;  APPLICANT: Bhatia, Ajay
```

```
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C8
;  CURRENT APPLICATION NUMBER: US/09/841,132
;  CURRENT FILING DATE:  2001-04-23
;  NUMBER OF SEQ ID NOS: 599
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 324
;   LENGTH: 1896
;   TYPE: DNA
;   ORGANISM: Chlamydia trachomatis
US-09-841-132-324

  Query Match             1.4%;  Score 41.4;  DB 10;  Length 1896;
  Best Local Similarity   54.2%;  Pred. No. 0.33;
  Matches   84;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy     869 TAGATGTTACAGGCAATCGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATG 928
           || |  ||  |||||||   |||   ||  |||||  | ||||  |  |||||      |
Db    1370 TAAAGTTTGAAGGCAATAAAGGTTCTATTGTTTTTGATTACAACTTTGCAAAAGGCAGAG 1429

Qy     929 GCGGAGCTATTTACGCTCCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAA 988
           |||||   ||      |    | | |  | || || |   |||   |      |||||  |
Db    1430 GCGGAAGCATCCTAACGAAAGAATTCTCTCTTGTAGCAGATGATTCGGTTGTCTTTAGTA 1489

Qy     989 ACAATATCGCCAATAATAAGGGGGGCGCTATCTAT 1023
           ||||||  ||  | ||    || || ||||| |||
Db    1490 ACAATACAGCAGAAAAAGGCGGTGGAGCTATTTAT 1524


RESULT 14
US-10-007-693-80
; Sequence 80, Application US/10007693
; Patent No. US20020146776A1
; GENERAL INFORMATION:
;  APPLICANT: Bhatia, Ajay
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
;  TITLE OF INVENTION:  AND DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.515C2
;  CURRENT APPLICATION NUMBER: US/10/007,693
;  CURRENT FILING DATE:  2001-12-05
;  NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 80
;   LENGTH: 3048
;   TYPE: DNA
;   ORGANISM: Chlamydia trachomatis serovar D
US-10-007-693-80

  Query Match             1.4%;  Score 41.4;  DB 12;  Length 3048;
  Best Local Similarity   54.2%;  Pred. No. 0.43;
  Matches   84;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy     869 TAGATGTTACAGGCAATCGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATG 928
           || |  ||  |||||||   |||   ||  |||||  | ||||  |  |||||      |
Db     929 TAAAGTTTGAAGGCAATAAAGGTTCTATTGTTTTTGATTACAACTTTGCAAAAGGCAGAG 988

Qy     929 GCGGAGCTATTTACGCTCCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAA 988
           |||||   ||      |    | | |  | || || |   |||   |      |||||  |
Db     989 GCGGAAGCATCCTAACGAAAGAATTCTCTCTTGTAGCAGATGATTCGGTTGTCTTTAGTA 1048

Qy     989 ACAATATCGCCAATAATAAGGGGGGCGCTATCTAT 1023
           ||||||  ||  | ||    || || ||||| |||
Db    1049 ACAATACAGCAGAAAAAGGCGGTGGAGCTATTTAT 1083


RESULT 15
US-09-815-242-8052
; Sequence 8052, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8052
;   LENGTH: 966
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(966)
US-09-815-242-8052

  Query Match             1.4%;  Score 40.8;  DB 10;  Length 966;
  Best Local Similarity   44.7%;  Pred. No. 0.33;
  Matches  159;  Conservative    0;  Mismatches  197;  Indels    0;  Gaps    0;

Qy    1223 ATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAA 1282
           || | |   ||| ||| || ||  |   |  || | |  ||| |    ||   | |  ||
Db     302 ATGCAAGACAAGATAGAAAAGCCCGTAGCCGTGAGCCAATCACTGCTAAATTAGTTGCAA 361

Qy    1283 CAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCTGCAGATTTTCATCAACGCAATT 1342
             | |  |  |  | | ||  | | | | |   ||| || |||  ||   |
Db     362 ACTTAATCGAAACAGCTGGCGCAACTCGTATGATTGCGTTAGACTTACATGCACCACAAA 421

Qy    1343 TACAAACAAAAACACCTGCACCCCTTACTCTCAGTAATGGTTTTCTATGTATCGAAGATC 1402
           | |||  |        |   ||  ||   | |   |  ||| | | |  | | |   | |
Db     422 TTCAAGGATTCTTTGATATTCCAATTGACCACTTAATGGGTGTGCCAATTCTTGCTAAAC 481

Qy    1403 ATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGGGGTGTTGTTTCTCTTGGGAATG 1462
           ||    |   |     ||||     | || ||    | |  |||||||| |  |   |||
Db     482 ATTTCAAAGATGATCCGAATATTAACCCAGAAGAATGTGTCGTTGTTTCACCAGACCATG 541

Qy    1463 GAGCAGTTCTGAGTTGCTATAAAAATGGTGCAGGAAATTCTGCTAGCAATGCCTCTATAA 1522
           | |  |||    ||     ||||   | ||            ||   | |||   ||||
Db     542 GCGGTGTTACACGTGCACGTAAATTAGCTGACATTTTAAAAACTCCAATTGCAATTATAG 601

Qy    1523 CACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTGAAAAGTGGTGCTGAGATT 1578
             | |  | ||   ||| || ||    |  |||| | || || |||||||
Db     602 ATAAACGTCGTCCTAGACCAAATGTTGCTGAAGTGATGAACATTGTTGGTGAGATT 657


Search completed: February  4, 2003, 20:53:39
Job time : 189 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 4, 2003, 15:57:07 ; Search time 3573 Seconds
(without alignments)
13135.886 Million cell updates/sec

Title: US-09-677-752-1
Perfect score: 2898
Sequence: 1 atgaaaaaagcgtttttctt.....aaattgctctgcgattctag 2898

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	49.6	1.7	1101	17	CNS0039G	AL063921 Drosophil
	2	44.6	1.5	667	13	BJ390296	BJ390296 BJ390296
c	3	44.6	1.5	806	17	CNS04AEE	AL281759 Tetraodon
	4	44	1.5	594	10	AW109947	AW109947 MT16306 m
c	5	44	1.5	1011	17	CNS0205Y	AL175183 Tetraodon
c	6	43.8	1.5	531	13	BJ438828	BJ438828 BJ438828
	7	43.8	1.5	597	9	AU261972	AU261972 AU261972
c	8	43.2	1.5	730	17	BH597226	BH597226 BOHPJ06TF
c	9	43	1.5	968	17	CNS006R5	AL065820 Drosophil
	10	42.8	1.5	600	14	BQ595967	BQ595967 PfESToab2
	11	41.6	1.4	538	17	AZ068468	AZ068468 RPCI-23-4
	12	41.4	1.4	1101	17	CNS0182P	AL108811 Drosophil
	13	41.2	1.4	667	12	BF298817	BF298817 021PbG02
c	14	41	1.4	312	13	BJ438184	BJ438184 BJ438184
c	15	41	1.4	415	13	BJ355703	BJ355703 BJ355703
	16	41	1.4	613	13	BJ425106	BJ425106 BJ425106
c	17	41	1.4	1101	17	CNS017DP	AL107911 Drosophil
	18	40.6	1.4	522	13	BI815912	BI815912 PfESToaa3
c	19	40.6	1.4	602	17	DR7O24T	AL735588 Danio rer
c	20	40.6	1.4	891	17	AZ529319	AZ529319 ENTBK28TR
c	21	40.6	1.4	919	17	AZ669216	AZ669216 ENTHP96TR
c	22	40.4	1.4	575	17	AQ981798	AQ981798 RPCI-23-3
c	23	40.2	1.4	487	17	AZ900612	AZ900612 RPCI-24-1
	24	40.2	1.4	562	17	PT013K11R	AL446148 Parameciu
c	25	40	1.4	493	17	AA550487	AA550487 1642m3 gm
c	26	39.8	1.4	324	14	F07305	F07305 HSC21G011 n
	27	39.8	1.4	350	9	AU060746	AU060746 AU060746
	28	39.8	1.4	388	9	AI831292	AI831292 wj82a02.x
	29	39.8	1.4	420	9	AJ280396	AJ280396 4A3A-AAO-
c	30	39.8	1.4	483	13	BJ405337	BJ405337 BJ405337
c	31	39.8	1.4	679	14	BM728778	BM728778 UI-E-EO1-
	32	39.6	1.4	409	10	BE580247	BE580247 kq38h11.y
	33	39.6	1.4	603	17	AQ439611	AQ439611 HS_5121_A
	34	39.6	1.4	1101	17	CNS00GDR	AL072060 Drosophil
	35	39.4	1.4	306	9	AA094179	AA094179 c12060.se
	36	39.4	1.4	526	13	BI814808	BI814808 PfESToaa0
	37	39.4	1.4	596	17	AQ803690	AQ803690 HS_3175_A
c	38	39.2	1.4	410	12	BE834190	BE834190 RC1-OT008
	39	39.2	1.4	451	9	AU268963	AU268963 AU268963
	40	39.2	1.4	643	13	BJ422941	BJ422941 BJ422941
	41	39.2	1.4	685	12	BF276133	BF276133 GA__EB002
c	42	39.2	1.4	723	13	BJ434535	BJ434535 BJ434535
	43	39.2	1.4	801	14	C24322	C24322 C24322 Dict
	44	39.2	1.4	894	17	AZ529017	AZ529017 ENTCV44TF
c	45	39.2	1.4	1074	17	CNS07E9K	AL441486 T3 end of

ALIGNMENTS

RESULT 1
CNS0039G/c
LOCUS CNS0039G 1101 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL063921
VERSION AL063921.1 GI:4941778
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial

EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACR08K10"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
BASE COUNT      201 a     64 c    131 g    202 t    503 others
ORIGIN
```

Query Match 1.7%; Score 49.6; DB 17; Length 1101;
Best Local Similarity 16.7%; Pred. No. 0.023;
Matches 88; Conservative 238; Mismatches 198; Indels 4; Gaps 2;

```
Qy   1064 ATGCTATTATTTTTAATGAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTA 1123
          :|: :::::|:::|:::::|:: :: |: : :|::: |  ::    ::::  :|  : ::
Db   1088 DTRKDDWDWTKWWTWWKDRADDRRWAGDADRWAWDDGAGTWWTATWWWWWWWWATWDTWW 1029

Qy   1124 CGTCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTC 1183
          :: :  |||:  |  |: :  ::|:   :::   : :: ||         :: |: :::
Db   1028 DKWWWWATAAKTDTAWTWWRTAWRADWAGRDRGAGKRDRDAATDADGAGRRDGGRKRKDK 969

Qy   1184 TATTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATG 1243
          :: ::|:::  ||: ::|   | :||::::::::::::::::  :: : :::: ||  :::
Db    968 KDRKDGDDDKKGGKKKKAAKAAKWATKWWDDWDWDKDWKWDGAKDRKADDDDGAGDKDDD 909

Qy   1244 CAGGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAG 1303
            |:: : : :||::  ::  ::::::| | :|:  :|: | : : :|:|: :::|: |
Db    908 GKGKDADDDTDGTKDDDDKDKWDDWDKAKGTWGDATWAWAATDWWWWGWADADWWTWDAA 849

Qy   1304 GAGCTACTGTTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCAC 1363
          :: :| ::::|: : :: | |:    |: :: | :|  ::  |: ::::  |  :: |
Db    848 ADDWWADDRWDAWAWKWDDAWAWGARTADRRDWGDRA-GKRGGARKRRDRKRADDKRDAA 790

Qy   1364 CCCTTACTCTCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGA--- 1420
             ::| | | : |  |::|::: : |:|:|  ||::| : : : ::     || ::
Db    789 DDRDDAATWTTWTTTTTRDTDDWKWKTDTWTRWAADRTWDRDDDDDRDRAGTAGRKWRRT 730

Qy   1421 ATCGATTCACACAAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCT 1480
          :: :::  : : ::|  ::  :::  :  :   : |:|  ||: : :::  | ::|:  :
Db    729 WKRRWKRRDTRWDDADADDTARDDRRRRGDDGADAGKGKKTGRKRRRRDRATWDRTDAWW 670

Qy   1481 ATAAAAATGGTGCAGGAAATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGAT 1540
          |:||:: | :|: :: : ::   :  :: ::|     ::::    :::|: :: |  :::
Db    669 ADAAWWTTTDTDTDDWDKRDRRRKGARRRRRTTARAAWDWWTWKAWDWAKWDWKTRADRW 610

Qy   1541 TGAATCTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGT 1588
          :::|  |:|     : :  :|:|:: ::::   |:|:   ||: :| |
Db    609 DRWAADTWTDARKADRDWAKARAWRARRDRARAARADRRWTTKGKTTT 562
```

```
RESULT 2
BJ390296
LOCUS       BJ390296                667 bp    mRNA    linear   EST 08-MAR-2002
DEFINITION  BJ390296 Dictyostelium discoideum cDNA library, SF Dictyostelium
            discoideum cDNA clone dds21123 5', mRNA sequence.
ACCESSION   BJ390296
VERSION     BJ390296.1  GI:19301382
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 667)
  AUTHORS   Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
  TITLE     Full length cDNA of Dictyostelium discoideum at the slug stage
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .667
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="dds21123"
                     /clone_lib="Dictyostelium discoideum cDNA library, SF"
                     /sex="mat A"
                     /dev_stage="Slug stage"
BASE COUNT      318 a     57 c     89 g    198 t      5 others
ORIGIN
```

Query Match 1.5%; Score 44.6; DB 13; Length 667;
Best Local Similarity 50.2%; Pred. No. 0.5;
Matches 110; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

```
Qy    931 GGAGCTATTTACGCTCCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAAC 990
          |||  | |  ||||||    | ||     | |||| |||       | | |   ||  ||
Db    257 GGAAATTTAAACGCTCACCCAATTGGTTCAATGGAAAATAATAACAACAATAATAATAAT 316

Qy    991 AATATCGCCAATAATAAGGGGGGCGCTATCTATATAGACGGAACCAGCAACTCCAAAATT 1050
          ||||     ||||||||         ||   |||| |    |  |  ||   ||| | |
Db    317 AATAATAATAATAATAATAATAATAATAATAATTTAATAATAATAATAATAACAATAAT 376

Qy   1051 TCTGCCGACCGCCATGCTATTATTTTTAATGAAAATATTGTGACTAATGTAACTAATGCA 1110
            |  | ||| | ||  || || |  |||| | |||| |    | ||||   | ||||
Db    377 AATAGCCACCACAATAATAATAATGGTAATAATAATAATAATAATAATAATAATAATAAT 436

Qy   1111 AATGGTACCAGTACGTCAGCTAATCCTCCTAGAAGAAAT 1149
          |||  ||  | ||       ||||  |  ||  || |||
Db    437 AATAATAGTAATAATAATAGTAATAATAATAATAGTAAT 475
```

```
RESULT 3
CNS04AEE/c
LOCUS       CNS04AEE                806 bp    DNA     linear   GSS 21-MAY-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence T7 end of clone
            095G24 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL281759
VERSION     AL281759.1  GI:8020086
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1  (bases 1 to 806)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 806)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
  TITLE     Charaterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 806)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000)
COMMENT     This sequence is a single read and was generated as part of a large
```

scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon.

FEATURES Location/Qualifiers
source 1. .806
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="095G24"
/clone_lib="G"
/note="Genoscope sequence ID : C0BG095BD12LP1-end : T7"

BASE COUNT 247 a 87 c 99 g 333 t 40 others
ORIGIN

Query Match 1.5%; Score 44.6; DB 17; Length 806;
Best Local Similarity 42.9%; Pred. No. 0.52;
Matches 106; Conservative 20; Mismatches 121; Indels 0; Gaps 0;

```
Qy    903 TAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTAGT 962
          || | |  |||  | :|| |||  | |     |||| |     |  |  :  || :  :
Db    672 TAATAATGATAATAAMAATAATAGTAGTACTACTACTACTAATAATAAMAATAAMAAMAA 613

Qy    963 GGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATCTA 1022
          | ||||     | : |   ||  || ||||     ||||||||  ||       || :
Db    612 MAAMAATGATAAAAAMAATAATAATAATAATAATAATAATAATGGTAATAGTAAMSS 553

Qy   1023 TATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAATGA 1082
          |||  | :  ||:| ::||  ::: :   :  :| :|: ::||  || || |  |||| |
Db    552 TATTAATVATACMAAHMACMAMMMCMSCCMCMMCCMCMAYMATAATAATAATAATAATAA 493

Qy   1083 AAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCTAG 1142
           |||| ||  | ||||   | |  |   |||  |   | ||       ||||  |  ||
Db    492 TAATAATGGTAATAATAATAATGGTAATAATAATGGTATTAATGATAGTAATGATAATAA 433

Qy   1143 AAGAAAT 1149
          || |||
Db    432 TAGTAAT 426
```

RESULT 4
AW109947
LOCUS AW109947 594 bp mRNA linear EST 31-JAN-2000
DEFINITION MT16306 mouse liver, dioxin treated Mus musculus cDNA clone MT16306 3', mRNA sequence.
ACCESSION AW109947
VERSION AW109947.1 GI:6822660
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 594)
AUTHORS Thomas,R.S., Rank,D.R., Penn,S.G., Zastrow,G.M., Jovanovich,S. and Bradfield,C.A.
TITLE McArdle Laboratory/Molecular Dynamics Dioxin EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Bradfield CA
McArdle Laboratory for Cancer Research
University of Wisconsin
1400 University Ave., Madison, WI 53706, USA
Tel: 608 262 2024
Fax: 608 262 2824
Email: bradfield@oncology.wisc.edu
This clone was sequenced as part of a project to develop a database on gene expression changes following exposure to various environmental toxicants. The database can be accessed at http://mcardle.oncology.wisc.edu/bradfield/. Treatment- Dioxin (10 ug/kg, ip) in a dioxane vehicle. Animals sacrified- 48 hrs post-injection.
Seq primer: dT(23)V (anchored polyT primer).

FEATURES Location/Qualifiers
source 1. .594
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="MT16306"
/clone_lib="mouse liver, dioxin treated"
/sex="male"
/lab_host="JM109"
/note="Organ: liver; Vector: pGEM11Zf (Promega); Site_1: NotI; Site_2: EcoRI; First strand cDNA was primed with a NotI-polyT primer [5'-AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTT-3']. Double-stranded cDNA was ligated with EcoRI adapters (Pharmacia), digested with NotI, and ligated into the EcoRI/NotI sites of the pGEM11Zf vector. The library was NOT normalized."

BASE COUNT 156 a 136 c 130 g 172 t
ORIGIN

Query Match 1.5%; Score 44; DB 10; Length 594;
Best Local Similarity 53.5%; Pred. No. 0.7;
Matches 92; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```
Qy    552 TAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTCTTTATGGACAA 611
          || || |  |||||    ||   ||  |    | || |||| |||| | ||  | || ||
Db     71 TAATATCTATAATATAAGTGACTTGGCCTTTGAACATTCTTTTTTTTTTTCCTTGAAAA 130

Qy    612 CATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCTGGAACGAGCAA 671
          ||||  | || ||| |     |  |||||||||   |  |     |  ||||  |   | |
Db    131 CATCATTTTTAAAAGTGCGTAATAAGGAAAAGGCCACTGTGCTGCTCATGGAGCCCACGA 190

Qy    672 TTCTTTTGAGAGTAATAACTGCGATCTCTTCTTTATCAATAACGCCTGTTGT 723
          | |        ||    ||||  |||||||| ||  |   ||| | || ||
Db    191 TCCCCACTTCCGTTCAGCCTGCAGTCTCTTCTCTAGGACATACGTCAGTGGT 242
```

RESULT 5
CNS0205Y/c
LOCUS CNS0205Y 1011 bp DNA linear GSS 12-MAY-2000
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone 221J14 of library G from Tetraodon nigroviridis, genomic survey sequence.
ACCESSION AL175183
VERSION AL175183.1 GI:7813240
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 1011)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.
TITLE Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1011)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
TITLE Charaterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1011)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000)
COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon.

FEATURES Location/Qualifiers
source 1. .1011
/organism="Tetraodon nigroviridis"

```
                     /db_xref="taxon:99883"
                     /clone="221J14"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0AG221DE07LP1-end : T7"
BASE COUNT       316 a    216 c    176 g    301 t      2 others
ORIGIN

  Query Match             1.5%;  Score 44;  DB 17;  Length 1011;
  Best Local Similarity   52.1%;  Pred. No. 0.82;
  Matches   98;  Conservative    0;  Mismatches   90;  Indels    0;  Gaps    0;

Qy     263 ATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTCACCCCTGAAAGTGGTGGTGCGA 322
           ||| | ||  | || ||| ||||||| | |   | ||   | ||| |||      || ||
Db     715 ATAAAAAATCATAATGTAATTATTTTTCCAGTGAACTTTTCTCTGGAAGCAACACTGGGA 656

Qy     323 TTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATTCGTGATACAATAGGTCCTGTAA 382
           ||  || || ||      ||  |    | ||||||  |  ||  || |   | ||
Db     655 TTAATTTTGAGAAGTATGATGATATCCCTGTGGAGGCTACTGGAACGAACTGCCCATCCC 596

Qy     383 TCTTTGAAAATAATACTTGTTGCAGACCATTTACATCGAGTAATCCTAATGCAGCTGTTA 442
           | |||||| |   | |   ||||   | |||  |  || || | || |       | |
Db     595 ACATTGAAAGTGTAAGTATATGCATTGCTTTTGAAATGATTACTACTCAAATTCAGGCTC 536

Qy     443 ATAAAATA 450
           ||||||| 
Db     535 TTAAAATA 528


RESULT 6
BJ438828/c
LOCUS       BJ438828                531 bp    mRNA    linear   EST 13-MAR-2002
DEFINITION  BJ438828 Dictyostelium discoideum cDNA library, VF Dictyostelium
            discoideum cDNA clone ddv38i15 3', mRNA sequence.
ACCESSION   BJ438828
VERSION     BJ438828.1  GI:19413550
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 531)
  AUTHORS   Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
  TITLE     Full length cDNA of Dictyostelium discoideum at the vegetative
            stage
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .531
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="ddv38i15"
                     /clone_lib="Dictyostelium discoideum cDNA library, VF"
                     /sex="mat A"
                     /dev_stage="Growth phase"
BASE COUNT       161 a     70 c     56 g    236 t      8 others
ORIGIN

  Query Match             1.5%;  Score 43.8;  DB 13;  Length 531;
  Best Local Similarity   45.1%;  Pred. No. 0.77;
  Matches  153;  Conservative    0;  Mismatches  186;  Indels    0;  Gaps    0;

Qy     258 TTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTCACCCCTGAAAGTGGTGG 317
           |||||||    |  | || | || |  |  |  | |  ||   |    |   |||    |
Db     436 TTTTGATGATGATGATGATGATAGTGGTCATATAGATGATGAAAATAATAGCAGTACAAG 377

Qy     318 TGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATTCGTGATACAATAGGTCC 377
           ||       | ||   | | | |   |   |      | | |  ||  |  |   |

Db     376 TGGTGAAAATGATATTAATACTACCAATATNAATAATAATAATAATGTCAATAGCAATAA 317

Qy     378 TGTAATCTTTGAAAATAATACTTGTTGCAGACCATTTACATCGAGTAATCCTAATGCAGC 437
           ||  | |  | | |||||||    | | |   |    ||  | | ||||  ||||
Db     316 TGATACCAATAATAATAATAACAATAGTAATACTACAACTACTAATAATAATAATAACAA 257

Qy     438 TGTTAATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAATCTTTACATAAATCATAA 497
           || || ||| |    | |           | | ||  | | |||  | | |  |||  |||
Db     256 TGATAGTAATAATTCAAACTCTANTAATAATAATAATAATAATAATAATAATAATANTAA 197

Qy     498 TCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCGAGGAGGAGCCATTAGTAC 557
           | ||||||     | |  ||  || | |   | |  |||  |  |    | || || ||
Db     196 TAATGATGATAATGAAAGTAATAATANCGGATATGGTGTAAGTAGTAATGGCAATAATAT 137

Qy     558 CGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTT 596
             || ||   ||| ||  || |  | |||   || |||
Db     136 TAATAGTAGTGTTGGTGGAAGTGTTACTCATCATTTTTT 98


RESULT 7
AU261972
LOCUS       AU261972                597 bp    mRNA    linear   EST 09-MAY-2002
DEFINITION  AU261972 VS Dictyostelium discoideum cDNA clone VSB358 3', mRNA
            sequence.
ACCESSION   AU261972
VERSION     AU261972.1  GI:20518519
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 597)
  AUTHORS   Urushihara,H., Morio,T., Saito,T., Koriki,E., Ochiai,H., Maeda,M.,
            Takeuchi,I., Kohara,Y. and Tanaka,Y.
  TITLE     Population analysis of cDNAs from unicellular and multicellular
            stages of Dictyostelium discoideum
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
            Tel: 81-298-53-4664
            Fax: 81-298-53-6614
            Email: hideko@biol.tsukuba.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .597
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="VSB358"
                     /clone_lib="VS"
                     /sex="mat A"
                     /dev_stage="vegetative"
BASE COUNT       336 a     59 c     51 g    151 t
ORIGIN

  Query Match             1.5%;  Score 43.8;  DB 9;  Length 597;
  Best Local Similarity   45.9%;  Pred. No. 0.8;
  Matches  150;  Conservative    0;  Mismatches  177;  Indels    0;  Gaps    0;

Qy     903 TAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTAGT 962
           || | | ||||| |  || ||| || |      || | |   |  |     ||   |
Db       8 TAATAATAATATTAATAATAATAATAGTAATAATAATAATAATAATAATAATAATTAA 67

Qy     963 GGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATCTA 1022
           |||||     ||   |  ||  || | |||    ||||||||         ||   |
Db      68 CAATAATAATAATAGTAATAATAGTAATAGTATTCATAATAATAATAATAATAGTAATAA 127

Qy    1023 TATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAATGA 1082
           ||   |    |  |  ||    || | |  |    |     ||  || || |  |||| |
Db     128 TAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 187

Qy    1083 AAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCTAG 1142
            |||||| | | || |  | ||||  |||  ||  | | |    | ||||  | |
```

```
Db    188 TAATATTATTAATAGTAATAGTAATAATAATAATAATTCAAATGGTAATATTACACC 247

Qy   1143 AAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGTAG 1202
          || ||   || |||||      |      |  | ||||| ||  || |   |  | |
Db    248 AACAACACCAGTAACAAGTAGTCGTAGAAGTATTCAAATTTTACAAGAAAATGCAAATCA 307

Qy   1203 CCAAAATTTAATTTTTTATGATCCTAT 1229
             || || |    || |  |||  ||
Db    308 TTTAATTTCACAAATTCAAAATCAAAT 334
```

RESULT 8
BH597226/c

LOCUS	BH597226 730 bp DNA linear GSS 15-DEC-2001
DEFINITION	BOHPJ06TF BOHP Brassica oleracea genomic clone BOHPJ06, DNA sequence.
ACCESSION	BH597226
VERSION	BH597226.1 GI:17849678
KEYWORDS	GSS.
SOURCE	Brassica oleracea.
ORGANISM	Brassica oleracea Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE	1 (bases 1 to 730)
AUTHORS	Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
TITLE	Whole genome shotgun sequencing of Brassica oleracea
JOURNAL	Unpublished (2001)
COMMENT	Other_GSSs: BOHPJ06TR Contact: Chris Town TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA. Tel: 301-838-3523 Fax: 301-838-0208 Email: cdtown@tigr.org DNA is from a doubled haploid provided by Tom Osborn. Seq primer: TF Class: sheared ends.

FEATURES	Location/Qualifiers
source	1. .730 /organism="Brassica oleracea" /strain="TO1000DH3" /db_xref="taxon:3712" /clone="BOHPJ06" /clone_lib="BOHP" /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOS1 using BstXI linkers"

BASE COUNT 225 a 115 c 128 g 262 t
ORIGIN

Query Match 1.5%; Score 43.2; DB 17; Length 730;
Best Local Similarity 47.4%; Pred. No. 1.2;
Matches 129; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

```
Qy    243 AGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTCAC 302
          | | |||   |  |  || || | | |||    | ||| | ||||  |||||||  |
Db    318 ACAATACTACATATAATTAGAAAGAAAAACTTTATGTAATAATTTACTAAAAAATAGCTA 259

Qy    303 CCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATTCG 362
            | |||  |       | | || |   || | | ||   ||  | |  | | | |
Db    258 AAGTAAAATAGTAAAAAAAAATTGTCAAAAGACTTCAAAAGTTCAAAACTAGCAAAGTAA 199

Qy    363 TGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACCATTTACATCGAG 422
          || ||   | |      ||||  |||   || || | |  | |   ||| || | ||  |
Db    198 TGCTATTTTGGTATTCTTAATTTTTTTTAAAAAAAATATTCGTGAACCTTTAATATTAAT 139

Qy    423 TAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAATCT 482
          || |  | |  || |   | ||   |   |   | | |   | |||    ||  ||
Db    138 TAGTTTTGTTTTAGGTTGGACTAGGTTGTTACTTGTCATTTCAATTGTAAATCTGAAAAA 79

Qy    483 TTACATAAATCATAATCATGATGTGGTCGGAT 514
          | | |||||| | |||   ||   || |||
Db     78 TAATATAAATTAAAATTTGTATAAATTCCGAT 47
```

RESULT 9
CNS006R5/c

LOCUS	CNS006R5 968 bp DNA linear GSS 03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR14N19 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION	AL065820
VERSION	AL065820.1 GI:4944788
KEYWORDS	GSS.
SOURCE	Drosophila melanogaster.
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE	1 (bases 1 to 968)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES	Location/Qualifiers
source	1. .968 /organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone="BACR14N19" /clone_lib="RPCI-98" /note="end : TET3"

BASE COUNT 249 a 280 c 60 g 197 t 182 others
ORIGIN

Query Match 1.5%; Score 43; DB 17; Length 968;
Best Local Similarity 42.5%; Pred. No. 1.5;
Matches 77; Conservative 16; Mismatches 88; Indels 0; Gaps 0;

```
Qy    903 TAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTAGT 962
          |||: :  | || |  |||| : :   ::   | ||   : |  |:|  |||   || |
Db    551 TAGWWRTKAAATNATKAAAAGGKGKATKKKTTTTTTGKKKKTTATRTNTTTANNTTATT 492

Qy    963 GGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATCTA 1022
            ||| |     |   |  | | || | |:|||     |  |||  ||  |   ||| |:
Db    491 WTATATTNTNTNTTTGTTNTGTTTANATAWTATATTATAANATANAGGATGNNATATWTW 432

Qy   1023 TATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAATGA 1082
           ||| | : |   | :    ||:::|| ||| :|  || | || || |  |  || | |
Db    431 AATAAATRAAGTGATMGNAGCCVMMATCTCTMMCCTCCCCNATNCTCTCTTGGTTTAGGT 372

Qy   1083 A 1083
          |
Db    371 A 371
```

RESULT 10
BQ595967

LOCUS	BQ595967 600 bp mRNA linear EST 24-JUN-2002
DEFINITION	PfESToab27d11.y1 Plasmodium falciparum 3D7 asexual cDNA Plasmodium falciparum cDNA 5', mRNA sequence.

```
ACCESSION   BQ595967
VERSION     BQ595967.1  GI:21542693
KEYWORDS    EST.
SOURCE      malaria parasite P. falciparum.
  ORGANISM  Plasmodium falciparum
            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE   1  (bases 1 to 600)
  AUTHORS   Tang,K., Cole,R., Chakrabarti,D., Haywood,R., Clifton,S., Pape,D.,
            Marra,M., Hillier,L., Martin,J., Wylie,T., Dante,M., Theising,B.,
            Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Jentes,E., Ronko,I.,
            Tsagareishvili,R., Belaygorod,L., Franklin,C., Carr,L., Grow,A.,
            Maguire,L., Richey,J., Wadkins,J., Kennedy,S., Levinso,D.,
            Waterston,R., Wilson,R. and Sibley,D.
  TITLE     WashU Plasmodium EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: L. David Sibley
            WashU Plasmodium EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Library was constructed by Debopam Chakrabarti DNA sequencing by:
            Washington University Genome Sequencing Center For information on
            obtaining a clone please contact: L. David Sibley
            (sibley@borcim.wustl.edu), Washington University
            Seq primer: -40UP from Gibco
            High quality sequence stop: 425.
FEATURES             Location/Qualifiers
     source          1. .600
                     /organism="Plasmodium falciparum"
                     /db_xref="taxon:5833"
                     /clone_lib="Plasmodium falciparum 3D7 asexual cDNA"
                     /lab_host="DH10B (GeneHog, Invitrogen, Inc.)"
                     /note="Vector: pBluescript SK plus; Site_1: EcoRI; Site_2:
                     XhoI; Library was constructed by Debopam Chakrabarti.
                     Total RNA samples were isolated from mixed stage
                     saponin(0.1%)-lysed P. falciparum 3D7 infected
                     erythrocytes by the acidic guanidinium-phenol chloroform
                     method. The poly A+ RNA was isolated by the polyAT-Tract
                     mRNA isolation system (Promega, WI) using streptavidin
                     Magnesphere particles. Directional cDNA libraries were
                     constructed by oligo d(T) priming of poly(A)+ RNA (5mg)
                     into EcoRI and XhoI sites of l ZapII vector using the Zap
                     cDNA synthesis kit (Stragene, CA). The average size of the
                     cDNA inserts in the library was between 1.0 and 1.5kb.
                     Clones were mass excised using the ExAssist helper phage
                     (Stratagene), the phagemids were precitptated with PEG
                     8000 and extracted with phenol/chloroform. Phagemid DNA
                     was electroporated into DH10B cells."
BASE COUNT      291 a     57 c     70 g    182 t
ORIGIN
```

```
  Query Match             1.5%;  Score 42.8;  DB 14;  Length 600;
  Best Local Similarity   47.1%;  Pred. No. 1.5;
  Matches  131;  Conservative    0;  Mismatches  147;  Indels    0;  Gaps    0;

Qy     965 ATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATCTATA 1024
           ||||||    ||   ||  ||| |  |||| |   |  |||||          |   |||
Db     219 ATAATGTGAATAATAACAATATGATTAATAACAATATGAATAATAACAATATGAATAATA 278

Qy    1025 TAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAATGAAA 1084
              |    | |  |||  ||  | |  |     |     ||  ||  | | | ||| | |
Db     279 ACAATATGAATAATAACAACATGAATAATATTAATGTGAATAATAGCAATATGAATAATA 338

Qy    1085 ATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCTAGAA 1144
           | |||   | |||  | | ||||   |||  |   | | |       || || || | |
Db     339 ACATTAATAATAACATGAATAATAACAATAATGTAAATGCTAGTAATATTCTTCTTGGTG 398

Qy    1145 GAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGTAGCC 1204
             ||||  ||||||  |||     |   ||  ||| || |   ||       | ||  |
Db     399 CTTCTGCATTAACAGGTGCAGCTATCAGCGGACAAAATCAAAATGGTATTAACAATAATC 458
```

```
Qy    1205 AAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAAT 1242
           |||||  |  ||  | || || | | | | |  |  | |
Db     459 AAAATGTTGTTAATAATAATACAAATAATGGAACTATT 496


RESULT 11
AZ068468
LOCUS       AZ068468                538 bp    DNA     linear   GSS 30-MAR-2000
DEFINITION  RPCI-23-420L3.TJ RPCI-23 Mus musculus genomic clone RPCI-23-420L3,
            DNA sequence.
ACCESSION   AZ068468
VERSION     AZ068468.1  GI:7359720
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 538)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 420  row: L  column: 3
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .538
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-420L3"
                     /clone_lib="RPCI-23"
                     /sex="Female"
                     /lab_host="DH10B"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies).  "
BASE COUNT      173 a     82 c    121 g    162 t
ORIGIN
```

```
  Query Match             1.4%;  Score 41.6;  DB 17;  Length 538;
  Best Local Similarity   59.2%;  Pred. No. 3.2;
  Matches   71;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy     804 TAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGATGGAGGAGCAATTAAAGTAACTAC 863
           | || || || ||  |   ||| || | |||| ||  ||||| |||  || || |||| |
Db     415 TTAAGATTTAAAACTAAAGTCTCCATATGCTTTTGTGGGAGGTGCAGCTACAGGAACTTC 474

Qy     864 TCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAA 923
           |  ||||| | |   ||  ||   |  | | | | ||  |   ||||||  ||| || |
Db     475 TGCCCTAGTTTTGCAAGAAAAAAATTATGAGCTTTCTTAGAACAACAATAATACAGAATA 534


RESULT 12
CNS0182P
```

```
LOCUS       CNS0182P                1101 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN37D10 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL108811
VERSION     AL108811.1  GI:5629115
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACN37D10"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : SP6"
BASE COUNT      274 a    268 c    128 g     73 t    358 others
ORIGIN

  Query Match             1.4%;  Score 41.4;  DB 17;  Length 1101;
  Best Local Similarity   14.6%;  Pred. No. 4.5;
  Matches   51;  Conservative  161;  Mismatches  137;  Indels    0;  Gaps    0;

Qy     365 ATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACCATTTACATCGAGTA 424
           ||| : ::  :  | :||| |: ::||| ::::   :| :  |:  :|  : :  | : |
Db     749 ATAWWTWDATWADTWKAATDTDAKRAAAGRRKRDARKTARDGGRRARTRRRAWAAGGKRA 808

Qy     425 ATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAATCTTT 484
           :    :|   ::  :: :::|:|| || | ||          ::: ::: : ::::: |
Db     809 RAGARRARAARRAADDRDWDAWAAAAAAAAAAAAAAWTTWRDRWWDWDDWDTRWDDDTTAA 868

Qy     485 ACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCGAGGAG 544
           : ::|:|  : ::  : :: : ::  ::| :: :  :::|  ||: :| : :  ::::::
Db     869 WWDDARARARRRRRRRRRRRRARRAARRAADDTDTKDRWADATTDKDTTKTWTTDDDDWD 928

Qy     545 GAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTCTTTA 604
           :|:  :::| :|    :::::  :::  | ::::   : :  :|:   ::::  :  :|
Db     929 KAKRDRWWAAKADGAWKWRDRARDDWAATAKDDDGWKDKWGGRGKRGGKDKKRWDKGKTG 988

Qy     605 TGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCTGGAA 664
           :::: ::  | |::::: : : ::::: :: :|: :: ::  |    |: :| : |:| :
Db     989 KKDDDDWDKTWTRDWWWWWTRTKWDWWWDDGRGRGRWTRKRWGAWWRADAWARDDTDGKD 1048

Qy     665 CGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTTATCAATAA 713
             :    : : |:||::: ::::::: :: ::: : :: |:::  :::::
Db    1049 TRTADKDRKRTDTTKRDGDDWRKDRKDKKRRDKKGDDKTKKDATWDDDD 1097


RESULT 13
BF298817
LOCUS       BF298817                667 bp    mRNA    linear   EST 04-MAY-2001
DEFINITION  021PbG02 Pb cDNA #20, Charles Yowell and Jane Carlton Plasmodium
            berghei cDNA 5', mRNA sequence.
ACCESSION   BF298817
VERSION     BF298817.1  GI:13952043
KEYWORDS    EST.
SOURCE      Plasmodium berghei.
  ORGANISM  Plasmodium berghei
            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE   1  (bases 1 to 667)
  AUTHORS   Carlton,J.M.-R. and Dame,J.B.
  TITLE     The Plasmodium vivax and P. berghei gene sequence tag projects
  JOURNAL   Parasitol. Today (Regul. Ed.) 16 (10), 409 (2000)
COMMENT     Contact: Dame JB
            Department of Pathobiology, College of Veterinary Medicine
            University of Florida
            2015 SW 23rd Avenue, Bldg 1017, Gainseville, FL 32611-0880, USA
            Tel: 352 392 4700
            Fax: 352 392 9704
            Email: damej@mail.vetmed.ufl.edu
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .667
                     /organism="Plasmodium berghei"
                     /strain="ANKA clone 15cy1 (clone of the ANKA 8417 HP clone
                     )"
                     /db_xref="taxon:5821"
                     /clone_lib="Pb cDNA #20, Charles Yowell and Jane Carlton"
                     /dev_stage="asynchronous blood stage"
                     /lab_host="Swiss white mice"
                     /note="Vector: pBluescript II vector DNA, excised from
                     Lamda ZAP II.; Site_1: EcoRI; Site_2: XhoI; Total RNA was
                     extracted from asynchronous blood stage forms of the
                     cloned ANKA isolate of P. berghei grown in laboratory
                     Swiss white mice. Contaminating host white cells had
                     previously been removed using a novel biomagnetic bead
                     protocol (J. Carlton et al., manuscript in preparation).
                     PolyA+ RNA was extracted and reverse transcribed using an
                     oligo dT-XhoI primer. Second strand cDNA was prepared
                     using RNase H and DNA polymerase I. EcoR I adaptors were
                     ligated to the cDNA, and it was digested with  XhoI.
                     Fragments were size selected, and those between 1-5 kb
                     ligated into EcoRI /XhoI digested vector."
BASE COUNT      339 a     49 c     70 g    209 t
ORIGIN

  Query Match             1.4%;  Score 41.2;  DB 12;  Length 667;
  Best Local Similarity   48.3%;  Pred. No. 4.4;
  Matches  115;  Conservative    0;  Mismatches  123;  Indels    0;  Gaps    0;

Qy     263 ATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTCACCCCTGAAAGTGGTGGTGCGA 322
           ||  | || || | | ||| |  | | | |  || | | |  | || | |    || |
Db     149 ATGAAAAACAATATGATATAGAAATGGAAGAGGATATAAACTGTAAAGGCGCAAATGATA 208

Qy     323 TTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATTCGTGATACAATAGGTCCTGTAA 382
           ||  ||||   |||   |||| |  ||      | | | || |||  | | ||  || ||
Db     209 TTTTTTATAACAGTTATAATTATTGTAATAATAATAATAGTAATAGTAAATGTGATGAAA 268

Qy     383 TCTTTGAAAATAATACTTGTTGCAGACCATTTACATCGAGTAATCCTAATGCAGCTGTTA 442
             | ||||| ||||       || |    |   |   |   ||  ||    ||    |
Db     269 AAATAGAAAAAAATATAGTAGACAAAAACATAGAAAATAAGTATAATATAAAAGAATATG 328

Qy     443 ATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAATCTTTACATAAATCATAATCA 500
           |||||| ||   || |   |    ||| |  | || |    |  | ||| || |||
Db     329 ATAAAACAAATAAAAGTATATTATTTCCTATTGAAGAAGAATTTAAAAAGATTATTCA 386


RESULT 14
BJ438184/c
LOCUS       BJ438184                312 bp    mRNA    linear   EST 13-MAR-2002
DEFINITION  BJ438184 Dictyostelium discoideum cDNA library, VF Dictyostelium
            discoideum cDNA clone ddv36k13 3', mRNA sequence.
ACCESSION   BJ438184
VERSION     BJ438184.1  GI:19412906
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
```

```
REFERENCE   1  (bases 1 to 312)
  AUTHORS   Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
  TITLE     Full length cDNA of Dictyostelium discoideum at the vegetative
            stage
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .312
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="ddv36k13"
                     /clone_lib="Dictyostelium discoideum cDNA library, VF"
                     /sex="mat A"
                     /dev_stage="Growth phase"
BASE COUNT       86 a     37 c     27 g    147 t     15 others
ORIGIN

  Query Match             1.4%;  Score 41;  DB 13;  Length 312;
  Best Local Similarity   45.8%;  Pred. No. 4;
  Matches  116;  Conservative    0;  Mismatches  137;  Indels    0;  Gaps    0;

Qy     871 GATGTTACAGGCAATCGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATGGC 930
           |||| |   |   || |||||   ||   |  |    | |||| ||  | || |  ||
Db     269 GATGATGATGATGATAGTGGTCATATANATGATGAAAATAATAGCAGNACAAGTGGTGAA 210

Qy     931 GGAGCTATTTACGCTCCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAAC 990
             | |||| |  ||||     | |          |||||| |  || | |   |   | |
Db     209 AATGATATTAATACTCCCNATATAAATAATAATAATGNCNATAGCAATAATGATACC 150

Qy     991 AATATCGCCAATAATAAGGGGGGCGCTATCTATATAGACGGAACCAGCAACTCCAAAATT 1050
            |||     ||||| ||  |     |||     |     |     |  |  |||    | | |
Db     149 NATAATAATAATAACAATAGTAATACTACNACTNCTAATAATAATAATAACAATGATAGT 90

Qy    1051 TCTGCCGACCGCCATGCTATTATTTTTAATGAAAATATTGTGACTAATGTAACTAATGCA 1110
             |         |  | ||| || |  |||| | |||| |   | ||||   | |||||
Db      89 AATAATTCNAACTNTNCTAATAATAATAATAATAATAATAATAATAATACTAATAATGAT 30

Qy    1111 AATGGTACCAGTA 1123
            ||      ||||
Db      29 GATAANGAAAGTA 17


RESULT 15
BJ355703/c
LOCUS       BJ355703                 415 bp    mRNA    linear   EST 07-MAR-2002
DEFINITION  BJ355703 Dictyostelium discoideum cDNA library, AF Dictyostelium
            discoideum cDNA clone dda57o22 3', mRNA sequence.
ACCESSION   BJ355703
VERSION     BJ355703.1  GI:19255297
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 415)
  AUTHORS   Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
  TITLE     Full length cDNA of Dictyostelium discoideum at the aggregation
            stage
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .415
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="dda57o22"
                     /clone_lib="Dictyostelium discoideum cDNA library, AF"
                     /sex="mat A"
                     /dev_stage="Aggregation stage"
BASE COUNT      130 a     53 c     41 g    189 t      2 others
ORIGIN

  Query Match             1.4%;  Score 41;  DB 13;  Length 415;
  Best Local Similarity   44.8%;  Pred. No. 4.3;
  Matches  152;  Conservative    0;  Mismatches  187;  Indels    0;  Gaps    0;

Qy     258 TTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTCACCCCTGAAAGTGGTGG 317
           || ||||     |  | || | || |  |  |  | | |  ||   |    |   |||    |
Db     385 TTCTGATGATGATGATGATGATAGTGGTCATATAGATGATGAAAATAATAGCAGTACAAG 326

Qy     318 TGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATTCGTGATACAATAGGTCC 377
           ||       | ||   | | | |   |   |      | | |  ||  | |    |
Db     325 TGGTGAAAATGATATTAATACTACCAATATNAATAATAATAATAATGTCAATAGCAATAA 266

Qy     378 TGTAATCTTTGAAAATAATACTTGTTGCAGACCATTTACATCGAGTAATCCTAATGCAGC 437
           ||  | |  | | |||||||    | | |   |    ||  | | ||||  ||||
Db     265 TGATACCAATAATAATAATAACAATAGTAATACTACAACTACTAATAATAATAATAACAA 206

Qy     438 TGTTAATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAATCTTTACATAAATCATAA 497
           || || ||| |    | |          | | ||  | | |||  | | |  |||  |||
Db     205 TGATAGTAATAATTCAAACTNTACTAATAATAATAATAATAATAATAATAATACTAA 146

Qy     498 TCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCGAGGAGGAGCCATTAGTAC 557
           | ||||||     | |  ||  || | |   | |  |||  |  |    | || || ||
Db     145 TAATGATGATAATGAAAGTAATAATAGCGGATATGGTGTAAGTAGTAATGGCAATAATAT 86

Qy     558 CGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTT 596
           || ||   ||| ||  || |  | |||   || |||
Db      85 TAATAGTAGTGTTGGTGGAAGTGTTACTCATCATTTTTT 47


Search completed: February  4, 2003, 19:08:01
Job time : 3587 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 4, 2003, 15:52:32 ; Search time 7193 Seconds
(without alignments)
11725.279 Million cell updates/sec

Title: US-09-677-752-1
Perfect score: 2898
Sequence: 1 atgaaaaaagcgtttttctt.....aaattgctctgcgattctag 2898

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		2821.2	97.3	2895	6	AX155928	AX155928 Sequence
2		2821.2	97.3	2895	6	AX361778	AX361778 Sequence
3	c	2818	97.2	14177	1	AE001360	AE001360 Chlamydia
4		2777.8	95.9	2934	6	AX155940	AX155940 Sequence
5		2777.8	95.9	2934	6	AX361790	AX361790 Sequence
6	c	1599.4	55.2	14482	1	AE002293	AE002293 Chlamydia
7	c	1190.6	41.1	1634	6	AX338273	AX338273 Sequence
8		136.4	4.7	2805	6	AX349485	AX349485 Sequence
9	c	136.4	4.7	10746	1	AE002190	AE002190 Chlamydop
10		136.4	4.7	12083	1	AE001631	AE001631 Chlamydia
11		136.4	4.7	300650	1	AP002546	AP002546 Chlamydop
12	c	124.2	4.3	10236	1	AE002189	AE002189 Chlamydop
13		88	3.0	2817	6	AX349483	AX349483 Sequence
14		56.6	2.0	14482	1	AE002293	AE002293 Chlamydia
15		48.4	1.7	2949	6	AX155927	AX155927 Sequence
16		48.4	1.7	2949	6	AX361777	AX361777 Sequence
17		48.4	1.7	3021	6	AX155939	AX155939 Sequence
18		48.4	1.7	3021	6	AX361789	AX361789 Sequence
19	c	48.2	1.7	107289	2	AC116923	AC116923 Dictyoste
20	c	47.6	1.6	160759	2	AC117082	AC117082 Dictyoste
21	c	47.2	1.6	298283	3	AE003782	AE003782 Drosophil
22		47	1.6	12838	1	AE002338	AE002338 Chlamydia
23		46	1.6	35276	2	AC115611	AC115611 Dictyoste
24		46	1.6	86904	2	AC117177	AC117177 Dictyoste
25		45.8	1.6	3353	3	AF482386	AF482386 Dictyoste
26		45.2	1.6	1941	6	AX156073	AX156073 Sequence
27		45.2	1.6	1941	6	AX361923	AX361923 Sequence
28		45.2	1.6	14177	1	AE001360	AE001360 Chlamydia
29		45	1.6	193924	2	AC095241	AC095241 Rattus no
30	c	44.8	1.5	40156	2	AC100852	AC100852 Homo sapi
31	c	44.8	1.5	161245	9	AC009158	AC009158 Homo sapi
32		44.8	1.5	189250	9	AC093511	AC093511 Homo sapi
33		44.6	1.5	105470	2	AC116306	AC116306 Dictyoste
34		44.2	1.5	4000	1	AF242295	AF242295 Bacillus
35	c	44.2	1.5	20043	2	AC116986	AC116986 Dictyoste
36	c	44	1.5	43993	2	AC116965	AC116965 Dictyoste
37	c	44	1.5	125020	9	AF429315	AF429315 Homo sapi
38		44	1.5	165399	2	AL840627	AL840627 Danio rer
39	c	43.8	1.5	152075	2	AC034178	AC034178 Homo sapi
40		43.6	1.5	6619	6	AX300903	AX300903 Sequence
41	c	43.6	1.5	56956	2	AC117081	AC117081 Dictyoste
42		43.6	1.5	96035	2	AC105319	AC105319 Oryza sat
43		43.6	1.5	181514	2	AC095643	AC095643 Rattus no
44	c	43.4	1.5	192929	2	AC005505	AC005505 Plasmodiu
45	c	43.2	1.5	172748	3	AC018484	AC018484 Drosophil

ALIGNMENTS

RESULT 1
AX155928
LOCUS AX155928 2895 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 171 from Patent WO0140474.
ACCESSION AX155928
VERSION AX155928.1 GI:14537033
KEYWORDS .
SOURCE Chlamydia sp.
ORGANISM Chlamydia sp.
Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE 1 (bases 1 to 2895)
AUTHORS Probst,P., Bhatia,A., Skeiky,Y.A., Fling,S.P. and Scholler,J.
TITLE Compounds and methods for treatment and diagnosis of chlamydial infection
JOURNAL Patent: WO 0140474-A 171 07-JUN-2001;

```
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .2895
                     /organism="Chlamydia sp."
                     /db_xref="taxon:35827"
BASE COUNT      866 a    632 c    558 g    839 t
ORIGIN

  Query Match             97.3%;  Score 2821.2;  DB 6;  Length 2895;
  Best Local Similarity   98.6%;  Pred. No. 0;
  Matches 2857;  Conservative    0;  Mismatches   38;  Indels    3;  Gaps    1;

Qy       1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATTGGAAACTCCCTATCAGGACTAGCTAGAGAG 60
           |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db       1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATCGGAAACTCCCTATCAGGACTAGCTAGAGAG 60

Qy      61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120

Qy     121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180

Qy     181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240

Qy     241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300

Qy     301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360

Qy     361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACCATTTACATCG 420
           |||||||||||||||||||||||||||||||||||||||||||||||| |||||||| |
Db     361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACTATTTACATGG 420

Qy     421 AGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 480
           || |||||| |||| ||||   ||||||||||||||||||||||||||||||||||||||
Db     421 AGAAATCCTTATGCTGCTG---ATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 477

Qy     481 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCGA 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     478 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCAA 537

Qy     541 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     538 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 597

Qy     601 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     598 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 657

Qy     661 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTTATCAATAACGCCTGT 720
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     658 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTCATCAATAACGCCTGT 717

Qy     721 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     718 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 777

Qy     781 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     778 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 837

Qy     841 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     838 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 897

Qy     901 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     898 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 957

Qy     961 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     958 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1017

Qy    1021 TATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1080
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    1018 TATATAGACGGAACCAGTAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1077

Qy    1081 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1078 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1137

Qy    1141 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1138 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1197

Qy    1201 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1198 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 1257

Qy    1261 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1258 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 1317

Qy    1321 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1318 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 1377

Qy    1381 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1378 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 1437

Qy    1441 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTGCAGGAAAT 1500
           |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||| ||
Db    1438 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTACAGGAGAT 1497

Qy    1501 TCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTG 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1498 TCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTG 1557

Qy    1561 AAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTAT 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1558 AAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTAT 1617

Qy    1621 ACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGAC 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1618 ACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGAC 1677

Qy    1681 TATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTGTCATCACAGCCTATG 1740
           || ||||| |||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    1678 TACGGGAACTCTCCTTATGAATCCACAGATCTGACCCATGCTCTGTCATCACAGCCTATG 1737

Qy    1741 CTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGATATGGATTTTTCGGGA 1800
           |||||||||||||| ||||| |||||||||||  ||| ||  |||| ||||||||||||
Db    1738 CTATCTATTTCTGAAGCTAGCGATAACCAGCTACAATCAGAAAATATAGATTTTTCGGGA 1797

Qy    1801 CTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAA 1860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1798 CTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAA 1857

Qy    1861 GATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAAGCCAATAGATTCCAT 1920
           |||||||||||||||||||||||||||||||| |||||||||||||||||||||||| |||
Db    1858 GATCCAGAACCAGCATCTTCAGCAACAATCACTGATCCACAAAAAGCCAATAGATTTCAT 1917

Qy    1921 AGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGCCCGAAACACAGAAGT 1980
           |||||||| |||| || |||||||||| |||||||||||||||| ||||||||||||
Db    1918 AGAACCTTACTACTAACATGGCTTCCTGCCGGGTATGTTCCTAGCCCAAAACACAGAAGT 1977

Qy    1981 CCCCTCATAGCGAATACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAAT 2040
           |||||||||||| || ||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    1978 CCCCTCATAGCTAACACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAAT 2037

Qy    2041 AGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 2100
           |||||||| |||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    2038 AGTGCAGAGCTGACACCTAGTGGTCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 2097

Qy    2101 ATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGA 2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2098 ATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGA 2157

Qy    2161 TACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTGAAATTCAGTCAGACC 2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2158 TACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTGAAATTCAGTCAGACC 2217

Qy    2221 TACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGC 2280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2218 TACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGC 2277

Qy    2281 CAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTT 2340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2278 CAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTT 2337

Qy    2341 TACAGCTATGGAGACCATAACTGTCACCATTTCTATACCCAAGGAGAAAATCTAACATCT 2400
           |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    2338 TACAGCTATGGAGACCATAACTGTCACCATTTCTATACTCAAGGAGAAAATCTAACATCT 2397

Qy    2401 CAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTGTTTTTTTTGATCTCCCTATGAAA 2460
           |||||||||||||| |||||||||||||||||||| ||||||||||||||||||||||||
Db    2398 CAAGGGACGTTCCGCAGTCAAACGATGGGAGGTGCTGTCTTTTTTGATCTCCCTATGAAA 2457

Qy    2461 CCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTTAGGTGCTCTTGGTATTTATTCT 2520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2458 CCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTTAGGTGCTCTTGGTATTTATTCT 2517

Qy    2521 AGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCT 2580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2518 AGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCT 2577

Qy    2581 TTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCAAAGA 2640
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    2578 TTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCACAGA 2637

Qy    2641 CCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACCAGGG 2700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2638 CCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACCAGGG 2697

Qy    2701 ATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCG 2760
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    2698 ATCGCGACCCAGCTCCTAGCCAGTAAAGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCG 2757

Qy    2761 CGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTC 2820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2758 CGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTC 2817

Qy    2821 CATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAA 2880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2818 CATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAA 2877

Qy    2881 ATTGCTCTGCGATTCTAG 2898
           ||||||||||||||||||
Db    2878 ATTGCTCTGCGATTCTAG 2895


RESULT 2
AX361778
LOCUS       AX361778               2895 bp    DNA     linear   PAT 15-FEB-2002
DEFINITION  Sequence 171 from Patent WO0208267.
ACCESSION   AX361778
VERSION     AX361778.1  GI:18694318
KEYWORDS    .
SOURCE      Chlamydia sp.
  ORGANISM  Chlamydia sp.
            Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.

REFERENCE   1
  AUTHORS   Fling,S.P., Skeiky,Y.A., Probst,P. and Bhatia,A.
  TITLE     Compounds and methods for treatment and diagnosis of chlamydial
            infection
  JOURNAL   Patent: WO 0208267-A 171 31-JAN-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .2895
                     /organism="Chlamydia sp."
                     /db_xref="taxon:35827"
BASE COUNT      866 a    632 c    558 g    839 t
ORIGIN

  Query Match             97.3%;  Score 2821.2;  DB 6;  Length 2895;
  Best Local Similarity   98.6%;  Pred. No. 0;
  Matches 2857;  Conservative    0;  Mismatches   38;  Indels    3;  Gaps    1;

Qy       1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATTGGAAACTCCCTATCAGGACTAGCTAGAGAG 60
           |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db       1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATCGGAAACTCCCTATCAGGACTAGCTAGAGAG 60

Qy      61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120

Qy     121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180

Qy     181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240

Qy     241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300

Qy     301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360

Qy     361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACCATTTACATCG 420
           |||||||||||||||||||||||||||||||||||||||||||||| |||||||| |||
Db     361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACTATTTACATGG 420

Qy     421 AGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 480
           || |||||| |||| ||||  ||||||||||||||||||||||||||||||||||||||||
Db     421 AGAAATCCTTATGCTGCTG---ATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 477

Qy     481 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCGA 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db     478 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCAA 537

Qy     541 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     538 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 597

Qy     601 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     598 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 657

Qy     661 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTTATCAATAACGCCTGT 720
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     658 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTCATCAATAACGCCTGT 717

Qy     721 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     718 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 777

Qy     781 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     778 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 837
```

```
Qy    841 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    838 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 897

Qy    901 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    898 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 957

Qy    961 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    958 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1017

Qy   1021 TATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1080
          ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   1018 TATATAGACGGAACCAGTAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1077

Qy   1081 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1078 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1137

Qy   1141 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1138 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1197

Qy   1201 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1198 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 1257

Qy   1261 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1258 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 1317

Qy   1321 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1318 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 1377

Qy   1381 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1378 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 1437

Qy   1441 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTGCAGGAAAT 1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||| ||
Db   1438 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTACAGGAGAT 1497

Qy   1501 TCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTG 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1498 TCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTG 1557

Qy   1561 AAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTAT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1558 AAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTAT 1617

Qy   1621 ACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGAC 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1618 ACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGAC 1677

Qy   1681 TATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTGTCATCACAGCCTATG 1740
          || ||||| |||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   1678 TACGGGAACTCTCCTTATGAATCCACAGATCTGACCCATGCTCTGTCATCACAGCCTATG 1737

Qy   1741 CTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGATATGGATTTTTCGGGA 1800
          |||||||||||||| ||||| |||||||||||  ||| ||  |||| ||||||||||||
Db   1738 CTATCTATTTCTGAAGCTAGCGATAACCAGCTACAATCAGAAAATATAGATTTTTCGGGA 1797

Qy   1801 CTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1798 CTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAA 1857

Qy   1861 GATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAAGCCAATAGATTCCAT 1920
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||| |||
Db   1858 GATCCAGAACCAGCATCTTCAGCAACAATCACTGATCCACAAAAAGCCAATAGATTTCAT 1917

Qy   1921 AGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGCCCGAAACACAGAAGT 1980
          |||||||| |||| || |||||||||| |||||||||||||||| ||||||||||||
Db   1918 AGAACCTTACTACTAACATGGCTTCCTGCCGGGTATGTTCCTAGCCCAAAACACAGAAGT 1977

Qy   1981 CCCCTCATAGCGAATACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAAT 2040
          |||||||||| || ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1978 CCCCTCATAGCTAACACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAAT 2037

Qy   2041 AGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 2100
          |||||||| |||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   2038 AGTGCAGAGCTGACACCTAGTGGTCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 2097

Qy   2101 ATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGA 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2098 ATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGA 2157

Qy   2161 TACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTGAAATTCAGTCAGACC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2158 TACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTGAAATTCAGTCAGACC 2217

Qy   2221 TACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2218 TACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGC 2277

Qy   2281 CAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2278 CAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTT 2337

Qy   2341 TACAGCTATGGAGACCATAACTGTCACCATTTCTATACCCAAGGAGAAAATCTAACATCT 2400
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   2338 TACAGCTATGGAGACCATAACTGTCACCATTTCTATACTCAAGGAGAAAATCTAACATCT 2397

Qy   2401 CAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTGTTTTTTTTGATCTCCCTATGAAA 2460
          ||||||||||||| |||||||||||||||||||||||| ||||||||||||||||||||||
Db   2398 CAAGGGACGTTCCGCAGTCAAACGATGGGAGGTGCTGTCTTTTTTGATCTCCCTATGAAA 2457

Qy   2461 CCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2458 CCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCT 2517

Qy   2521 AGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCT 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2518 AGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCT 2577

Qy   2581 TTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCAAAGA 2640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   2578 TTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCACAGA 2637

Qy   2641 CCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACCAGGG 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2638 CCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACCAGGG 2697

Qy   2701 ATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCG 2760
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   2698 ATCGCGACCCAGCTCCTAGCCAGTAAAGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCG 2757

Qy   2761 CGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTC 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2758 CGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTC 2817

Qy   2821 CATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAA 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2818 CATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAA 2877

Qy   2881 ATTGCTCTGCGATTCTAG 2898
          ||||||||||||||||||
Db   2878 ATTGCTCTGCGATTCTAG 2895
```

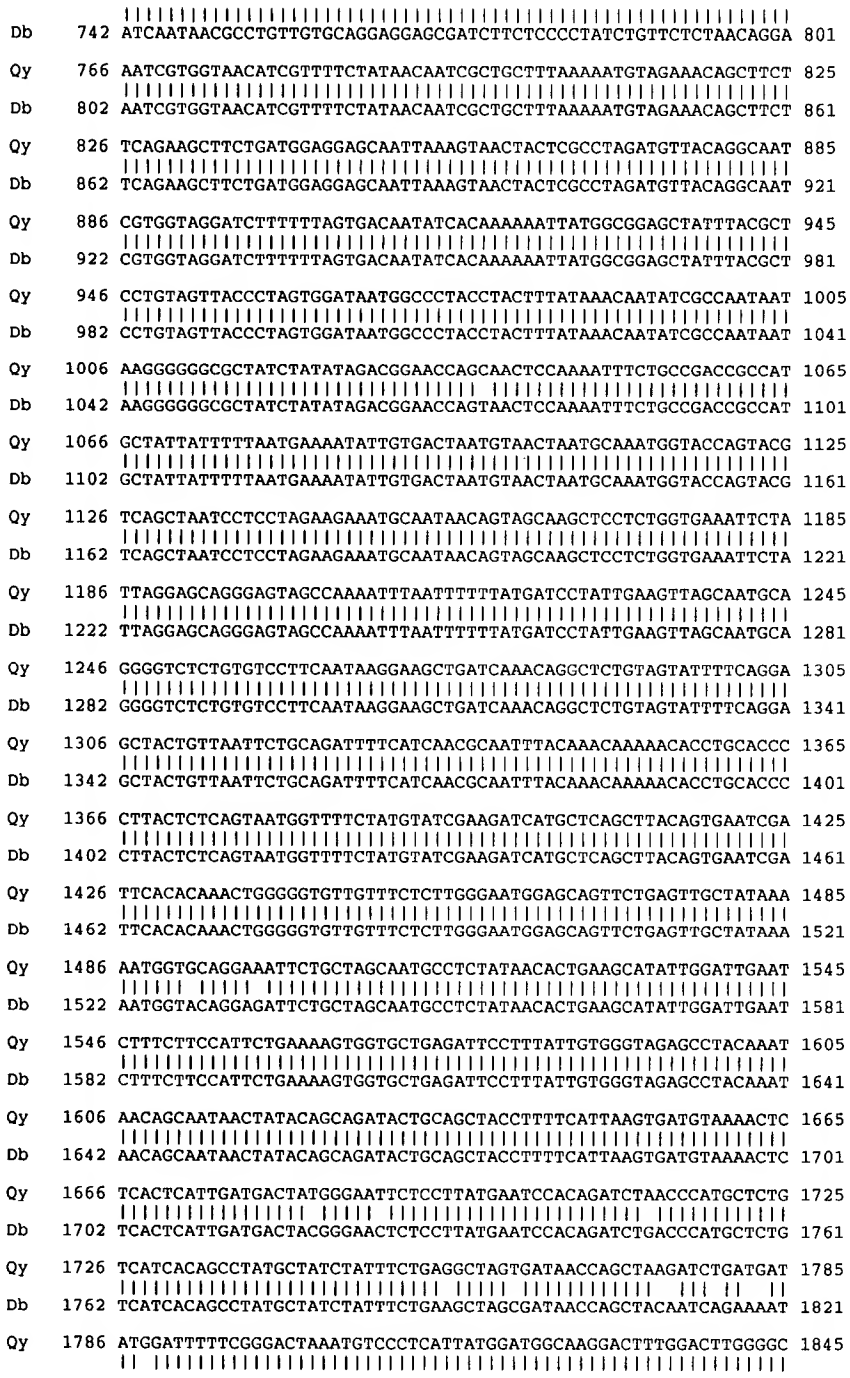
RESULT 3
AE001360/c
LOCUS AE001360 14177 bp DNA linear BCT 30-OCT-2000
DEFINITION Chlamydia trachomatis section 87 of 87 of the complete genome.
ACCESSION AE001360 AE001273

```
VERSION     AE001360.1  GI:3329342
KEYWORDS    .
SOURCE      Chlamydia trachomatis.
  ORGANISM  Chlamydia trachomatis
            Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE   1  (bases 1 to 14177)
  AUTHORS   Stephens,R.S., Kalman,S., Lammel,C.J., Fan,J., Marathe,R.,
            Aravind,L., Mitchell,W.P., Olinger,L., Tatusov,R.L., Zhao,Q.,
            Koonin,E.V. and Davis,R.W.
  TITLE     Genome sequence of an obligate intracellular pathogen of humans:
            Chlamydia trachomatis
  JOURNAL   Science 282 (5389), 754-759 (1998)
  MEDLINE   99000809
   PUBMED   9784136
REFERENCE   2  (bases 1 to 14177)
  AUTHORS   Kalman,S., Mitchell,W., Marathe,R., Lammel,C., Fan,J., Hyman,R.W.,
            Olinger,L., Grimwood,J., Davis,R.W. and Stephens,R.S.
  TITLE     Comparative genomes of Chlamydia pneumoniae and C. trachomatis
  JOURNAL   Nat. Genet. 21 (4), 385-389 (1999)
  MEDLINE   99206606
   PUBMED   10192388
REFERENCE   3  (bases 1 to 14177)
  AUTHORS   Stephens,R.S., Kalman,S., Lammel,C.J., Fan,J., Marathe,R.,
            Aravind,L., Mitchell,W.P., Olinger,L., Tatusov,R.L., Zhao,Q.,
            Koonin,E.V. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAY-1998) Program in Infectious Diseases, University
            of California, 235 Warren Hall, Berkeley, CA 94720-7360, USA
FEATURES             Location/Qualifiers
     source          1. .14177
                     /organism="Chlamydia trachomatis"
                     /strain="D/UW-3/CX"
                     /db_xref="taxon:813"
     gene            complement(137. .1393)
                     /gene="CT868"
     CDS             complement(137. .1393)
                     /gene="CT868"
                     /codon_start=1
                     /transl_table=11
                     /product="Membrane Thiol Protease (predicted)"
                     /protein_id="AAC68466.1"
                     /db_xref="GI:3329343"
                     /translation="MLSPTNSTSKKAPVPPQDSSKPVLISEEPQNQLLQKVARTALAV
                     LLVVVTLGLILLFYSFSDLQSFPWCCQTRPSTKEQPTISIPVPLPSPPLAVPRPSTPP
                     PPVISRPSTPPAPTPAISPPSTPSAPKPSTPPPLPPKAPKPVKTQEDLLPFVPEQVFV
                     EMYEDMARRWIIEALVPAWDSDIIFKCLCYFHTLYQGLIPLETFPPATIFNFKQKIIS
                     ILEDKKAVLRGEPIKGSLPICCSEENYRRHLHGTTLLPVFMWYHPTPKTLSDTMQTMK
                     QLAIKGSVGASHWLLVIVDIQARRLVYFDSLYNYVMSPEDMEKDLQSFAQQLDQVYPA
                     YDSQKFSVKIAAKEVIQKGSGSSCGAWCCQFLHWYLRDPFTDALNDLPVDSVERHENL
                     ASFVQACEAAVQDLPELFWPEAKALF"
     gene            complement(1570. .4464)
                     /gene="pmpE"
                     /note="CT869"
     CDS             complement(1570. .4464)
                     /gene="pmpE"
                     /codon_start=1
                     /transl_table=11
                     /product="Putative Outer Membrane Protein E"
                     /protein_id="AAC68467.1"
                     /db_xref="GI:3329344"
                     /translation="MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKI
                     SLTGDTHNLTNCYLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTP
                     ESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRLFTWRNPYAADKIREGGAIHAQNL
                     YINHNHDVVGFMKNFSYVQGGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIY
                     AGTSNSFESNNCDLFFINNACCAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSE
                     ASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANN
                     KGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGE
                     ILLGAGSSQNLIFYDPIEVSNAGVSVSFNKEADQTGSVVFSGATVNSADFHQRNLQTK
                     TPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGTGDSASNASITL
                     KHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYE
                     STDLTHALSSQPMLSISEASDNQLQSENIDFSGLNVPHYGWQGLWTWGWAKTQDPEPA
                     SSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAE
                     LTPSGHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFSLKFSQTY
                     TKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQGENLT
                     SQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSFST
                     KTPLINVLVPIGVKGSFMNATHRPQAWTVELAYQPVLYRQEPGIAAQLLASKGIWFGS
                     GSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF"
     gene            complement(4467. .7571)
                     /gene="pmpF"
                     /note="CT870"
     CDS             complement(4467. .7571)
                     /gene="pmpF"
                     /codon_start=1
                     /transl_table=11
                     /product="Putative Outer Membrane Protein F"
                     /protein_id="AAC68468.1"
                     /db_xref="GI:3329345"
                     /translation="MIKRTSLSFACLSFFYLSTISILQANETDTLQFRRFTFSDREIQ
                     FVLDPASLITAQNIVLSNLQSNGTGACTISGNTQTQIFSNSVNTTADSGGAFDMVTTS
                     FTASDNANLLFCNNYCTHNKGGGAIRSGGPIRFLNNQDVLFYNNISAGAKYVGTGDHN
                     EKNRGGALYATTITLTGNRTLAFINNMSGDCGGAISADTQISITDTVKGILFENNHTL
                     NHIPYTQAENMARGGAICSRRDLCSISNNSGPIVFNYNQGGKGGAISATRCVIDNNKE
                     RIIFSNNSSLGWSQSSSASNGGAIQTTQGFTLRNNKGSIYFDSNTATHAGGAINCGYI
                     DIRDNGPVYFLNNSAAWGAAFNLSKPRSATNYIHTGTGDIVFNNNVVFTLDGNLLGKR
                     KLFHINNNEITPYTLSLGAKKDTRIYFYDLFQWERVKENTSNNPPSPTSRNTITVNPE
                     TEFSGAVVFSYNQMSSDIRTLMGKEHNYIKEAPTTLKFGTLAIEDDAELEIFNIPFTQ
                     NPTSLLALGSGATLTVGKHGKLNITNLGVILPIILKEGKSPPCIRVNPQDMTQNTGTG
                     QTPSSTSSISTPMIIFNGRLSIVDENYESVYDSMDLSRGKAEQLILSIETTNDGQLDS
                     NWQSSLNTSLLSPPHYGYQGLWTPNWITTTYTITLNNNSSAPTSATSIAEQKKTSETF
                     TPSNTTTASIPNIKASAGSGSGSASNSGEVTITKHTLVVNWAPVGYIVDPIRRGDLIA
                     NSLVHSGRNMTMGLRSLLPDNSWFALQGAATTLFTKQQKRLSYHGYSSASKGYTVSSQ
                     ASGAHGHKFLLSFSQSSDKMKEKETNNRLSSRYYLSALCFEHPMFDRIALIGAAACNY
                     GTHNMRSFYGTKKSSKGKFHSTTLGASLRCELRDSMPLRSIMLTPFAQALFSRTEPAS
                     IRESGDLARLFTLEQAHTAVVSPIGIKGAYSSDTWPTLSWEMELAYQPTLYWKRPLLN
                     TLLIQNNGSWVTTNTPLAKHSFYGRGSHSLKFSHLKLFANYQAEVATSTVSHYINAGG
                     ALVF"
     gene            7736. .10777
                     /gene="pmpG"
                     /note="CT871"
     CDS             7736. .10777
                     /gene="pmpG"
                     /codon_start=1
                     /transl_table=11
                     /product="Putative Outer Membrane Protein G"
                     /protein_id="AAC68469.1"
                     /db_xref="GI:3329346"
                     /translation="MQTSFHKFFLSMILAYSCCSLSGGGYAAEIMIPQGIYDGETLTV
                     SFPYTVIGDPSGTTVFSAGELTLKNLDNSIAALPLSCFGNLLGSFTVLGRGHSLTFEN
                     IRTSTNGAALSDSANSGLFTIEGFKELSFSNCNSLLAVLPAATTNNGSQTPTTTSTPS
                     NGTIYSKTDLLLLNNEKFSFYSNLVSGDGGAIDAKSLTVQGISKLCVFQENTAQADGG
                     ACQVVTSFSAMANEAPIAFIANVAGVRGGGIAAVQDGQQGVSSSTSTEDPVVSFSRNT
                     AVEFDGNVARVGGGIYSYGNVAFLNNGKTLFLNNVASPVYIAAEQPTNGQASNTSDNY
                     GDGGAIFCKNGAQAAGSNNSGSVSFDGEGVVFFSSNVAAGKGGAIYAKKLSVANCGPV
                     QFLGNIANDGGAIYLGESGELSLSADYGDIIFDGNLKRTAKENAADVNGVTVSSQAIS
                     MGSGGKITTLRAKAGHQILFNDPIEMANGNNQPAQSSEPLKINDGEGYTGDIVFANGN
                     STLYQNVTIEQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTLDFVTPQPPQQPPAANQ
                     LITLSNLHLSLSSLLANNAVTNPPTNPPAQDSHPAIIGSTTAGSVTISGPIFFEDLDD
                     TAYDRYDWLGSNQKIDVLKLQLGTQPSANAPSDLTLGNEMPKYGYQGSWKLAWDPNTA
                     NNGPYTLKATWTKTGYNPGPERVASLVPNSLWGSILDIRSAHSAIQASVDGRSYCRGL
                     WVSGVSNFFYHDRDALGQGYRYISGGYSLGANSYFGSSMFGLAFTEVFGRSKDYVVCR
                     SNHHACIGSVYLSTKQALCGSYLFGDAFIRASYGFGNQHMKTSYTFAEESDVRWDNNC
                     LVGEIGVGLPIVITPSKLYLNELRPFVQAEFSYADHESFTEEGDQARAFRSGHLMNLS
                     VPVGVKFDRCSSTHPNKYSFMGAYICDAYRTISGTQTTLLSHQETWTTDAFHLARHGV
                     IVRGSMYASLTSNIEVYGHGRYEYRDTSRGYGLSAGSKVRF"
     gene            10808. .13858
                     /gene="pmpH"
                     /note="CT872"
     CDS             10808. .13858
                     /gene="pmpH"
                     /codon_start=1
                     /transl_table=11
                     /product="Putative Outer Membrane Protein H"
                     /protein_id="AAC68470.1"
                     /db_xref="GI:3329347"
                     /translation="MPFSLRSTSFCFLACLCSYSYGFASSPQVLTPNVTTPFKGDDVY
                     LNGDCAFVNVYAGAENGSIISANGDNLTITGQNHTLSFTDSQGPVLQNYAFISAGETL
```

```
                    TLKDFSSLMFSKNVSCGEKGMISGKTVSISGAGEVIFWDNSVGYSPLSIVPASTPTPP
                    APAPAPAASSSLSPTVSDARKGSIFSVETSLEISGVKKGVMFDNNAGNFGTVFRGNSN
                    NNAGSGGSGSATTPSFTVKNCKGKVSFTDNVASCGGGVVYKGTVLFKDNEGGIFFRGN
                    TAYDDLGILAATSRDQNTETGGGGGVICSPDDSVKFEGNKGSIVFDYNFAKGRGGSIL
                    TKEFSLVADDSVVFSNNTAEKGGGAIYAPTIDISTNGGSILFERNRAAEGGAICVSEA
                    SSGSTGNLTLSASDGDIVFSGNMTSDRPGERSAARILSDGTTVSLNASGLSKLIFYDP
                    VVQNNSAAGASTPSPSSSSMPGAVTINQSGNGSVIFTAESLTPSEKLQVLNSTSNFPG
                    ALTVSGGELVVTEGATLTTGTITATSGRVTLGSGASLSAVAGAANNNYTCTVSKLGID
                    LESFLTPNYKTAILGADGTVTVNSGSTLDLVMESEAEVYDNPLFVGSLTIPFVTLSSS
                    SASNGVTKNSVTINDADAAHYGYQGSWSADWTKPPLAPDAKGMVPPNTNNTLYLTWRP
                    ASNYGEYRLDPQRKGELVPNSLWVAGSALRTFTNGLKEHYVSRDVGFVASLHALGDYI
                    LNYTQDDRDGFLARYGGFQATAASHYENGSIFGVAFGQLYGQTKSRMYYSKDAGNMTM
                    LSCFGRSYVDIKGTETVMYWETAYGYSVHRMHTQYFNDKTQKFDHSKCHWHNNNYYAF
                    VGAEHNFLEYCIPTRQFARDYELTGFMRFEMAGGWSSSTRETGSLTRYFARGSGHNMS
                    LPIGIVAHAVSHVRRSPPSKLTLNMGYRPDIWRVTPHCNMEIIANGVKTPIQGSPLAR
                    HAFFLEVHDTLYIHHFGRAYMNYSLDARRRQTAHFVSMGLNRIF"
BASE COUNT     4016 a    2583 c    3328 g    4250 t
ORIGIN

  Query Match             97.2%;  Score 2818;  DB 1;  Length 14177;
  Best Local Similarity   98.5%;  Pred. No. 0;
  Matches 2855;  Conservative    0;  Mismatches   40;  Indels    3;  Gaps    1;

Qy       1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATTGGAAACTCCCTATCAGGACTAGCTAGAGAG 60
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    4464 ATGAAAAAAGCGTTTTTCTTTTTCCTTATCGGAAACTCCCTATCAGGACTAGCTAGAGAG 4405

Qy      61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4404 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 4345

Qy     121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4344 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 4285

Qy     181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4284 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 4225

Qy     241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4224 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 4165

Qy     301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4164 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 4105

Qy     361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACCATTTACATCG 420
          |||||||||||||||||||||||||||||||||||||||||||||||| |||||||| |
Db    4104 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACTATTTACATGG 4045

Qy     421 AGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 480
          || |||||| |||| ||||   ||||||||||||||||||||||||||||||||||||||
Db    4044 AGAAATCCTTATGCTGCTG---ATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 3988

Qy     481 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCGA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    3987 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCAA 3928

Qy     541 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3927 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 3868

Qy     601 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3867 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 3808

Qy     661 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTTATCAATAACGCCTGT 720
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    3807 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTCATCAATAACGCCTGT 3748

Qy     721 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    3747 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 3688

Qy     781 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3687 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 3628

Qy     841 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3627 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 3568

Qy     901 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3567 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 3508

Qy     961 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3507 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 3448

Qy    1021 TATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1080
          |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db    3447 TATATAGACGGAACCAGTAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 3388

Qy    1081 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3387 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 3328

Qy    1141 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3327 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 3268

Qy    1201 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3267 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 3208

Qy    1261 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3207 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 3148

Qy    1321 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3147 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 3088

Qy    1381 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3087 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 3028

Qy    1441 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTGCAGGAAAT 1500
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||| ||
Db    3027 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTACAGGAGAT 2968

Qy    1501 TCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTG 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2967 TCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTG 2908

Qy    1561 AAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTAT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2907 AAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTAT 2848

Qy    1621 ACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGAC 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2847 ACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGAC 2788

Qy    1681 TATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTGTCATCACAGCCTATG 1740
          || ||||| |||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    2787 TACGGGAACTCTCCTTATGAATCCACAGATCTGACCCATGCTCTGTCATCACAGCCTATG 2728

Qy    1741 CTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGATATGGATTTTTCGGGA 1800
          |||||||||||||| ||||| |||||||||||  ||| ||  |||| ||||||||||||
Db    2727 CTATCTATTTCTGAAGCTAGCGATAACCAGCTACAATCAGAAAATATAGATTTTTCGGGA 2668

Qy    1801 CTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2667 CTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAA 2608
```

```
Qy   1861 GATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAAGCCAATAGATTCCAT 1920
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||| |||
Db   2607 GATCCAGAACCAGCATCTTCAGCAACAATCACTGATCCACAAAAAGCCAATAGATTTCAT 2548

Qy   1921 AGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGCCCGAAACACAGAAGT 1980
          ||||||||| |||| || |||||||||| ||||||||||||||||| ||||||||||||
Db   2547 AGAACCTTACTACTAACATGGCTTCCTGCCGGGTATGTTCCTAGCCCAAAACACAGAAGT 2488

Qy   1981 CCCCTCATAGCGAATACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAAT 2040
          |||||||||| || ||||||||||||||||||||||||||||||||||||||||||||||
Db   2487 CCCCTCATAGCTAACACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAAT 2428

Qy   2041 AGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 2100
          |||||||| |||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   2427 AGTGCAGAGCTGACACCTAGTGGTCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGC 2368

Qy   2101 ATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGA 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2367 ATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGA 2308

Qy   2161 TACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTGAAATTCAGTCAGACC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2307 TACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTGAAATTCAGTCAGACC 2248

Qy   2221 TACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2247 TACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGC 2188

Qy   2281 CAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2187 CAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTT 2128

Qy   2341 TACAGCTATGGAGACCATAACTGTCACCATTTCTATACCCAAGGAGAAAATCTAACATCT 2400
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   2127 TACAGCTATGGAGACCATAACTGTCACCATTTCTATACTCAAGGAGAAAATCTAACATCT 2068

Qy   2401 CAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTGTTTTTTTTGATCTCCCTATGAAA 2460
          |||||||||||||| |||||||||||||||||||| ||||||||||||||||||||||||
Db   2067 CAAGGGACGTTCCGCAGTCAAACGATGGGAGGTGCTGTCTTTTTTGATCTCCCTATGAAA 2008

Qy   2461 CCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2007 CCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCT 1948

Qy   2521 AGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCT 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1947 AGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCT 1888

Qy   2581 TTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCAAAGA 2640
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   1887 TTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCACAGA 1828

Qy   2641 CCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACCAGGG 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1827 CCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACCAGGG 1768

Qy   2701 ATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCG 2760
          |||||  |||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   1767 ATCGCAGCCCAGCTCCTAGCCAGTAAGGGTATTTGGTTCGGTAGTGGAAGCCCCTCATCG 1708

Qy   2761 CGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTC 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1707 CGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTC 1648

Qy   2821 CATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAA 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1647 CATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAA 1588

Qy   2881 ATTGCTCTGCGATTCTAG 2898
          |||||||||||||||||
Db   1587 ATTGCTCTGCGATTCTAG 1570
```

```
RESULT 4
AX155940
LOCUS       AX155940               2934 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 183 from Patent WO0140474.
ACCESSION   AX155940
VERSION     AX155940.1  GI:14537039
KEYWORDS    .
SOURCE      Chlamydia sp.
  ORGANISM  Chlamydia sp.
            Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE   1  (bases 1 to 2934)
  AUTHORS   Probst,P., Bhatia,A., Skeiky,Y.A., Fling,S.P. and Scholler,J.
  TITLE     Compounds and methods for treatment and diagnosis of chlamydial
            infection
  JOURNAL   Patent: WO 0140474-A 183 07-JUN-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .2934
                     /organism="Chlamydia sp."
                     /db_xref="taxon:35827"
BASE COUNT      875 a    645 c    575 g    839 t
ORIGIN

  Query Match             95.9%;  Score 2777.8;  DB 6;  Length 2934;
  Best Local Similarity   98.6%;  Pred. No. 0;
  Matches 2813;  Conservative    0;  Mismatches   37;  Indels    3;  Gaps    1;

Qy     46 GGACTAGCTAGAGAGGTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCT 105
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     85 GGACTAGCTAGAGAGGTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCT 144

Qy    106 ACGAAAGAGTCGCTATCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAAC 165
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    145 ACGAAAGAGTCGCTATCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAAC 204

Qy    166 TGCTATCTCGATAACCTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGA 225
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    205 TGCTATCTCGATAACCTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGA 264

Qy    226 GCTGCTGTCACAATAACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTAT 285
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    265 GCTGCTGTCACAATAACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTAT 324

Qy    286 TTTGCAAAAAATCTCACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCT 345
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    325 TTTGCAAAAAATCTCACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCT 384

Qy    346 CCTACCGTGGAGATTCGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGC 405
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    385 CCTACCGTGGAGATTCGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGC 444

Qy    406 AGACCATTTACATCGAGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCC 465
          |||| |||||||| ||| |||||| |||| ||||   ||||||||||||||||||||||
Db    445 AGACTATTTACATGGAGAAATCCTTATGCTGCTG---ATAAAATAAGAGAAGGCGGAGCC 501

Qy    466 ATTCATGCTCAAAATCTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAAC 525
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    502 ATTCATGCTCAAAATCTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAAC 561

Qy    526 TTTTCTTATGTCCGAGGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAAT 585
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    562 TTTTCTTATGTCCAAGGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAAT 621

Qy    586 CAGTCTTGTTTTCTCTTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGT 645
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    622 CAGTCTTGTTTTCTCTTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGT 681

Qy    646 GGCGCTATCTATGCTGGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTT 705
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    682 GGCGCTATCTATGCTGGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTC 741

Qy    706 ATCAATAACGCCTGTTGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGA 765
```

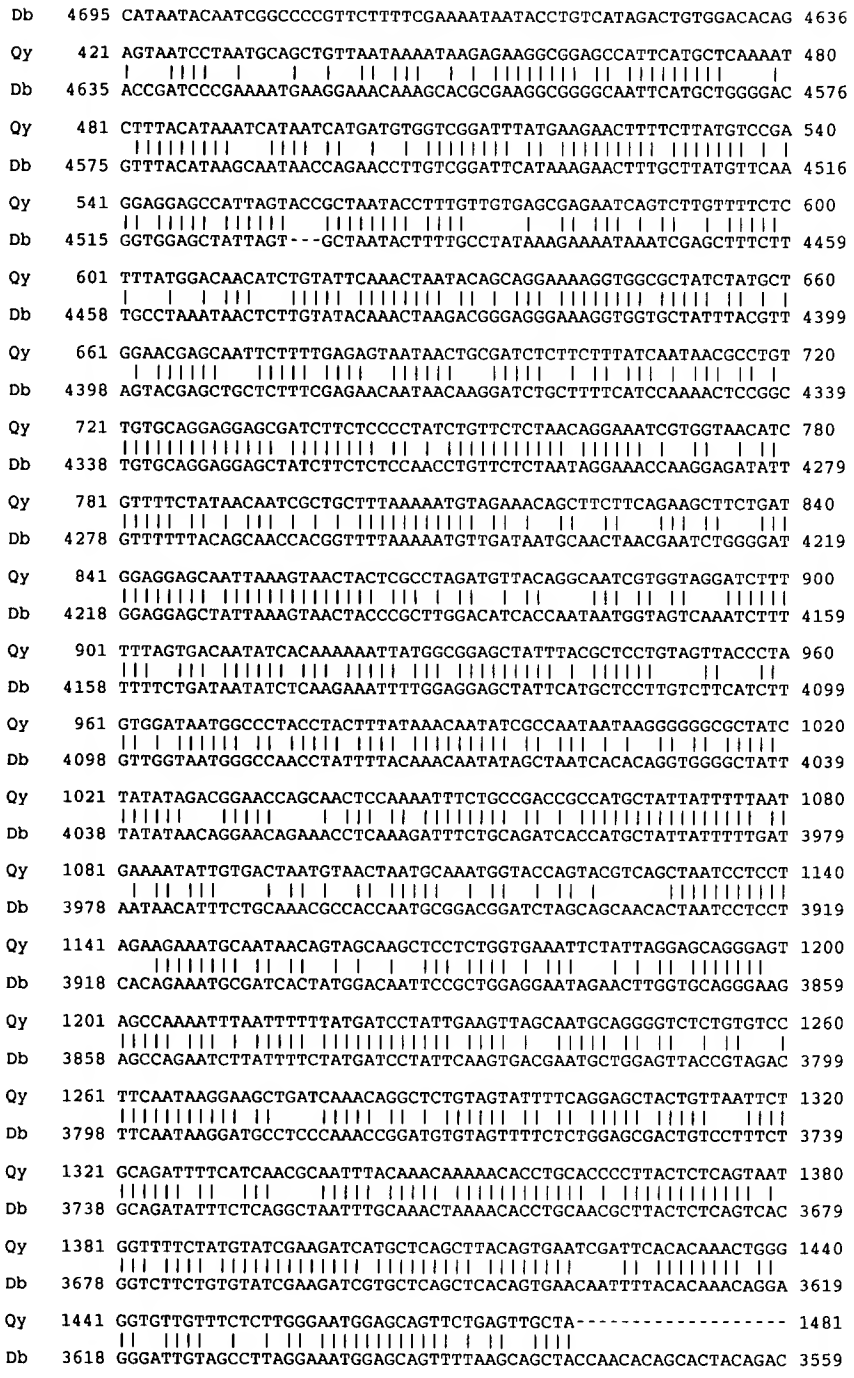

Db 742 ATCAATAACGCCTGTTGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGA 801
Qy 766 AATCGTGGTAACATCGTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCT 825
Db 802 AATCGTGGTAACATCGTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCT 861
Qy 826 TCAGAAGCTTCTGATGGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAAT 885
Db 862 TCAGAAGCTTCTGATGGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAAT 921
Qy 886 CGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCT 945
Db 922 CGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCT 981
Qy 946 CCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAAT 1005
Db 982 CCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAAT 1041
Qy 1006 AAGGGGGGCGCTATCTATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCAT 1065
Db 1042 AAGGGGGGCGCTATCTATATAGACGGAACCAGTAACTCCAAAATTTCTGCCGACCGCCAT 1101
Qy 1066 GCTATTATTTTTAATGAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACG 1125
Db 1102 GCTATTATTTTTAATGAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACG 1161
Qy 1126 TCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTA 1185
Db 1162 TCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTA 1221
Qy 1186 TTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCA 1245
Db 1222 TTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCA 1281
Qy 1246 GGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGA 1305
Db 1282 GGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGA 1341
Qy 1306 GCTACTGTTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCC 1365
Db 1342 GCTACTGTTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCC 1401
Qy 1366 CTTACTCTCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGA 1425
Db 1402 CTTACTCTCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGA 1461
Qy 1426 TTCACACAAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAA 1485
Db 1462 TTCACACAAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAA 1521
Qy 1486 AATGGTGCAGGAAATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAAT 1545
Db 1522 AATGGTACAGGAGATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAAT 1581
Qy 1546 CTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAAT 1605
Db 1582 CTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAAT 1641
Qy 1606 AACAGCAATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTC 1665
Db 1642 AACAGCAATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTC 1701
Qy 1666 TCACTCATTGATGACTATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTG 1725
Db 1702 TCACTCATTGATGACTACGGGAACTCTCCTTATGAATCCACAGATCTGACCCATGCTCTG 1761
Qy 1726 TCATCACAGCCTATGCTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGAT 1785
Db 1762 TCATCACAGCCTATGCTATCTATTTCTGAAGCTAGCGATAACCAGCTACAATCAGAAAAT 1821
Qy 1786 ATGGATTTTTCGGGACTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGC 1845
Db 1822 ATAGATTTTTCGGGACTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGC 1881
Qy 1846 TGGGCAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAA 1905
Db 1882 TGGGCAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACTGATCCACAAAAA 1941
Qy 1906 GCCAATAGATTCCATAGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGC 1965
Db 1942 GCCAATAGATTTCATAGAACCTTACTACTAACATGGCTTCCTGCCGGGTATGTTCCTAGC 2001
Qy 1966 CCGAAACACAGAAGTCCCCTCATAGCGAATACCTTATGGGGGAATATGCTGCTTGCAACA 2025
Db 2002 CCAAAACACAGAAGTCCCCTCATAGCTAACACCTTATGGGGGAATATGCTGCTTGCAACA 2061
Qy 2026 GAAAGCTTAAAAAATAGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTACA 2085
Db 2062 GAAAGCTTAAAAAATAGTGCAGAGCTGACACCTAGTGGTCATCCTTTCTGGGGAATTACA 2121
Qy 2086 GGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCAT 2145
Db 2122 GGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCAT 2181
Qy 2146 ATGCGCTCTTCCGGATACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTG 2205
Db 2182 ATGCGCTCTTCCGGATACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTG 2241
Qy 2206 AAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCT 2265
Db 2242 AAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCT 2301
Qy 2266 AAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACT 2325
Db 2302 AAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACT 2361
Qy 2326 AAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCACCATTTCTATACCCAAGGA 2385
Db 2362 AAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCACCATTTCTATACTCAAGGA 2421
Qy 2386 GAAAATCTAACATCTCAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTGTTTTTTTT 2445
Db 2422 GAAAATCTAACATCTCAAGGGACGTTCCGCAGTCAAACGATGGGAGGTGCTGTCTTTTTT 2481
Qy 2446 GATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCT 2505
Db 2482 GATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCT 2541
Qy 2506 CTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTT 2565
Db 2542 CTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTT 2601
Qy 2566 TCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATG 2625
Db 2602 TCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATG 2661
Qy 2626 AATGCTACCCAAAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTAT 2685
Db 2662 AATGCTACCCACAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTAT 2721
Qy 2686 AGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGTTTGGTAGT 2745
Db 2722 AGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAGTAAAGGTATTTGGTTTGGTAGT 2781
Qy 2746 GGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTG 2805
Db 2782 GGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTG 2841
Qy 2806 AGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAAT 2865
Db 2842 AGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAAT 2901
Qy 2866 TATCTCAATGGGGAAATTGCTCTGCGATTCTAG 2898
Db 2902 TATCTCAATGGGGAAATTGCTCTGCGATTCTAG 2934

```
RESULT 5
AX361790
LOCUS       AX361790               2934 bp    DNA     linear   PAT 15-FEB-2002
DEFINITION  Sequence 183 from Patent WO0208267.
ACCESSION   AX361790
VERSION     AX361790.1  GI:18694324
KEYWORDS    .
SOURCE      Chlamydia sp.
  ORGANISM  Chlamydia sp.
            Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE   1
  AUTHORS   Fling,S.P., Skeiky,Y.A., Probst,P. and Bhatia,A.
  TITLE     Compounds and methods for treatment and diagnosis of chlamydial
            infection
  JOURNAL   Patent: WO 0208267-A 183 31-JAN-2002;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .2934
                     /organism="Chlamydia sp."
                     /db_xref="taxon:35827"
BASE COUNT      875 a    645 c    575 g    839 t
ORIGIN

  Query Match             95.9%;  Score 2777.8;  DB 6;  Length 2934;
  Best Local Similarity   98.6%;  Pred. No. 0;
  Matches 2813;  Conservative    0;  Mismatches   37;  Indels    3;  Gaps    1;

Qy      46 GGACTAGCTAGAGAGGTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCT 105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85 GGACTAGCTAGAGAGGTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCT 144

Qy     106 ACGAAAGAGTCGCTATCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAAC 165
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     145 ACGAAAGAGTCGCTATCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAAC 204

Qy     166 TGCTATCTCGATAACCTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGA 225
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     205 TGCTATCTCGATAACCTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGA 264

Qy     226 GCTGCTGTCACAATAACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTAT 285
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     265 GCTGCTGTCACAATAACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTAT 324

Qy     286 TTTGCAAAAAATCTCACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCT 345
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     325 TTTGCAAAAAATCTCACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCT 384

Qy     346 CCTACCGTGGAGATTCGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGC 405
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     385 CCTACCGTGGAGATTCGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGC 444

Qy     406 AGACCATTTACATCGAGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCC 465
           |||| |||||||| ||| |||||| |||| ||||   ||||||||||||||||||||||
Db     445 AGACTATTTACATGGAGAAATCCTTATGCTGCTG---ATAAAATAAGAGAAGGCGGAGCC 501

Qy     466 ATTCATGCTCAAAATCTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAAC 525
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     502 ATTCATGCTCAAAATCTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAAC 561

Qy     526 TTTTCTTATGTCCGAGGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAAT 585
           |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db     562 TTTTCTTATGTCCAAGGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAAT 621

Qy     586 CAGTCTTGTTTTCTCTTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGT 645
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     622 CAGTCTTGTTTTCTCTTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGT 681

Qy     646 GGCGCTATCTATGCTGGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTT 705
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     682 GGCGCTATCTATGCTGGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTC 741

Qy     706 ATCAATAACGCCTGTTGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGA 765
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     742 ATCAATAACGCCTGTTGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGA 801

Qy     766 AATCGTGGTAACATCGTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCT 825
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     802 AATCGTGGTAACATCGTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCT 861

Qy     826 TCAGAAGCTTCTGATGGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAAT 885
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     862 TCAGAAGCTTCTGATGGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAAT 921

Qy     886 CGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCT 945
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     922 CGTGGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCT 981

Qy     946 CCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAAT 1005
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     982 CCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAAT 1041

Qy    1006 AAGGGGGGCGCTATCTATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCAT 1065
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    1042 AAGGGGGGCGCTATCTATATAGACGGAACCAGTAACTCCAAAATTTCTGCCGACCGCCAT 1101

Qy    1066 GCTATTATTTTTAATGAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACG 1125
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1102 GCTATTATTTTTAATGAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACG 1161

Qy    1126 TCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTA 1185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1162 TCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTA 1221

Qy    1186 TTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCA 1245
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1222 TTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCA 1281

Qy    1246 GGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGA 1305
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1282 GGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGA 1341

Qy    1306 GCTACTGTTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCC 1365
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1342 GCTACTGTTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCC 1401

Qy    1366 CTTACTCTCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGA 1425
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1402 CTTACTCTCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGA 1461

Qy    1426 TTCACACAAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAA 1485
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1462 TTCACACAAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAA 1521

Qy    1486 AATGGTGCAGGAAATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAAT 1545
           |||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    1522 AATGGTACAGGAGATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAAT 1581

Qy    1546 CTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAAT 1605
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1582 CTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAAT 1641

Qy    1606 AACAGCAATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTC 1665
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1642 AACAGCAATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTC 1701

Qy    1666 TCACTCATTGATGACTATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTG 1725
           |||||||||||||||| ||||| ||||||||||||||||||||| ||||||||||||
Db    1702 TCACTCATTGATGACTACGGGAACTCTCCTTATGAATCCACAGATCTGACCCATGCTCTG 1761

Qy    1726 TCATCACAGCCTATGCTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGAT 1785
           |||||||||||||||||||||||||||| ||||| |||||||||||  ||| ||  ||
Db    1762 TCATCACAGCCTATGCTATCTATTTCTGAAGCTAGCGATAACCAGCTACAATCAGAAAAT 1821

Qy    1786 ATGGATTTTTCGGGACTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGC 1845
```

```
Db   1822 ATAGATTTTTCGGGACTAAATGTCCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGC 1881

Qy   1846 TGGGCAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAA 1905

Db   1882 TGGGCAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACTGATCCACAAAAA 1941

Qy   1906 GCCAATAGATTCCATAGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGC 1965

Db   1942 GCCAATAGATTTCATAGAACCTTACTACTAACATGGCTTCCTGCCGGGTATGTTCCTAGC 2001

Qy   1966 CCGAAACACAGAAGTCCCCTCATAGCGAATACCTTATGGGGGAATATGCTGCTTGCAACA 2025

Db   2002 CCAAAACACAGAAGTCCCCTCATAGCTAACACCTTATGGGGGAATATGCTGCTTGCAACA 2061

Qy   2026 GAAAGCTTAAAAAATAGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTACA 2085

Db   2062 GAAAGCTTAAAAAATAGTGCAGAGCTGACACCTAGTGGTCATCCTTTCTGGGGAATTACA 2121

Qy   2086 GGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCAT 2145

Db   2122 GGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCCAT 2181

Qy   2146 ATGCGCTCTTCCGGATACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTG 2205

Db   2182 ATGCGCTCTTCCGGATACTCTGCGGGGATGATAGCAGGGCAGACACACACCTTCTCATTG 2241

Qy   2206 AAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCT 2265

Db   2242 AAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCT 2301

Qy   2266 AAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACT 2325

Db   2302 AAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACT 2361

Qy   2326 AAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCACCATTTCTATACCCAAGGA 2385

Db   2362 AAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCACCATTTCTATACTCAAGGA 2421

Qy   2386 GAAAATCTAACATCTCAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTGTTTTTTTT 2445

Db   2422 GAAAATCTAACATCTCAAGGGACGTTCCGCAGTCAAACGATGGGAGGTGCTGTCTTTTTT 2481

Qy   2446 GATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCT 2505

Db   2482 GATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCT 2541

Qy   2506 CTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTT 2565

Db   2542 CTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTT 2601

Qy   2566 TCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATG 2625

Db   2602 TCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATG 2661

Qy   2626 AATGCTACCCAAAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTAT 2685

Db   2662 AATGCTACCCACAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTAT 2721

Qy   2686 AGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGTTTGGTAGT 2745

Db   2722 AGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAGTAAAGGTATTTGGTTTGGTAGT 2781

Qy   2746 GGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTG 2805

Db   2782 GGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTG 2841

Qy   2806 AGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAAT 2865

Db   2842 AGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAAT 2901

Qy   2866 TATCTCAATGGGGAAATTGCTCTGCGATTCTAG 2898
```

```
Db   2902 TATCTCAATGGGGAAATTGCTCTGCGATTCTAG 2934


RESULT 6
AE002293/c
LOCUS       AE002293               14482 bp    DNA     linear   BCT 26-MAY-2000
DEFINITION  Chlamydia muridarum, section 25 of 85 of the complete genome.
ACCESSION   AE002293 AE002160
VERSION     AE002293.1  GI:7190298
KEYWORDS    .
SOURCE      Chlamydia muridarum.
  ORGANISM  Chlamydia muridarum
            Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE   1  (bases 1 to 14482)
  AUTHORS   Read,T.D., Brunham,R., Shen,C., Gill,S.R., Heidelberg,J.F.,
            White,O., Hickey,E.K., Peterson,J., Umayam,L.A., Utterback,T.,
            Berry,K., Bass,S., Linher,K., Weidman,J., Khouri,H., Craven,B.,
            Bowman,C., Dodson,R., Gwinn,M., Nelson,W., DeBoy,R., Kolonay,J.,
            McClarty,G., Salzberg,S.L., Eisen,J. and Fraser,C.M.
  TITLE     Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
            pneumoniae AR39
  JOURNAL   Nucleic Acids Res. 28 (6), 1397-1406 (2000)
  MEDLINE   20150255
   PUBMED   10684935
REFERENCE   2  (bases 1 to 14482)
  AUTHORS   Read,T.D., Brunham,R., Shen,C., Gill,S.R., Heidelberg,J.F.,
            White,O., Hickey,E.K., Peterson,J., Umayam,L.A., Utterback,T.,
            Berry,K., Bass,S., Linher,K., Weidman,J., Khouri,H., Craven,B.,
            Bowman,C., Dodson,R., Gwinn,M., Nelson,W., DeBoy,R., Kolonay,J.,
            McClarty,G., Salzberg,S.L., Eisen,J. and Fraser,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAR-2000) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .14482
                     /organism="Chlamydia muridarum"
                     /db_xref="taxon:83560"
                     /note="synonym: Chlamydia trachomatis MoPn"
     gene            complement(355. .825)
                     /gene="TC0259"
     CDS             complement(355. .825)
                     /gene="TC0259"
                     /note="hypothetical protein; identified by Glimmer2;
                     putative"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAF39128.1"
                     /db_xref="GI:7190299"
                     /translation="MVSPSYPSAKTISASLHESPQQLLVQKRNRWSFLQKVVRIALAV
                     LVVIMTLGLILFFYTFSEITTFPWCCRKHLLPKEDIIYPLASPSPSLQHLPSIHEEKA
                     PAPKTPELLKAKPYISTYPLLESINTKESLVPIISEKVVLEFFEEISTRKRIQS"
     gene            complement(937. .2046)
                     /gene="TC0260"
     CDS             complement(937. .2046)
                     /gene="TC0260"
                     /note="hypothetical protein; identified by Glimmer2;
                     putative"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAF39129.1"
                     /db_xref="GI:7190300"
                     /translation="MAKISSFSSLPTFDAPKSPPSLFSQSSLLGKIARIIVALFLILV
                     SLGLILIAYRFSDLLNCKFCITKTTELPIPSTKIPISSPPVPTTPPYQKKEPTLPQKN
                     PKISENLLSQEVVQDYLSSGRLPELAILDNSQMFQFMCVLHDQYPKLLPNDCLIPLTI
                     FNYREEICNTIQDKLKADQGQYCSLGDLQCPITCSPENYHQLLQQSRVLPFLLWYDPE
                     PTNHQQTLEKMQEIASQGTPGNSHWTVIVVDLDARCITYFDSLVNYIASTDEMERRMK
                     SLACCLANIGLCKNNGCPFDVHVAVNESLQNWMGSCCGLWCCQYMKWYMDHSHTGILQ
                     KIPDSLAYKTLLLQSLHSTFEKLMKKYADLSWPTT"
     gene            complement(2125. .5055)
                     /gene="TC0261"
     CDS             complement(2125. .5055)
```

```
                     /gene="TC0261"
                     /note="similar to GP:4376753; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="polymorphic membrane protein E/F family"
                     /protein_id="AAF39130.1"
                     /db_xref="GI:7190301"
                     /translation="MKKLFFFVLIGSSILGFTREVPPSILLKPILNPYHMTGLFFPKV
                     NLLGDTHNLTDYHLDNLKCILACLQRTPYEGAAFTVTDYLGFSDTQKDGIFCFKNLTP
                     ESGGVIGSPTQNTPTIKIHNTIGPVLFENNTCHRLWTQTDPENEGNKAREGGAIHAGD
                     VYISNNQNLVGFIKNFAYVQGGAISANTFAYKENKSSFLCLNNSCIQTKTGGKGGAIY
                     VSTSCSFENNNKDLLFIQNSGCAGGAIFSPTCSLIGNQGDIVFYSNHGFKNVDNATNE
                     SGDGGAIKVTTRLDITNNGSQIFFSDNISRNFGGAIHAPCLHLVGNGPTYFTNNIANH
                     TGGAIYITGTETSKISADHHAIIFDNNISANATNADGSSSNTNPPHRNAITMDNSAGG
                     IELGAGKSQNLIFYDPIQVTNAGVTVDFNKDASQTGCVVFSGATVLSADISQANLQTK
                     TPATLTLSHGLLCIEDRAQLTVNNFTQTGGIVALGNGAVLSSYQHSTTDATQTPPTTT
                     TTDASVTLNHIGLNLPSILKDGAEMPLLWVEPISTTQGNTTTYTSDTAASFSLNGATL
                     SLIDEDGNSPYENTDLSRALYAQPMLAISEASDNQLQSESMDFSKVNVPHYGWQGLWT
                     WGWAKTENPTTTPPATITDPKKANQFHRTLLLTWLPAGYIPSPKHKSPLIANTLWGNI
                     LFATENLKNSSGQELLDRPFWGITGGGLGMMVYQEPRKDHPGFHMHTSGYSAGMITGN
                     THTFSLRFSQSYTKLNERYAKNYVSSKNYSCQGEMLLSLQEGLMLTKLIGLYSYGNHN
                     SHHFYTQGEDLSSQGEFHSQTFGGAVFFDLPLKPFGRTHILTAPFLGAIGMYSKLSSF
                     TEVGAYPRTFITETPLINVLIPIGVKGSFMNATHRPQAWTVELAYQPVLYRQEPSIST
                     QLLAGKGMWFGHGSPASRHALAYKISQKTQLLRFATLQLQYHGYYSSSTFCNYLNGEV
                     SLRF"
gene                 complement(5058. .8135)
                     /gene="TC0262"
CDS                  complement(5058. .8135)
                     /gene="TC0262"
                     /note="similar to GP:4376753; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="polymorphic membrane protein E/F family"
                     /protein_id="AAF39131.1"
                     /db_xref="GI:7190302"
                     /translation="MTRRILPLSLVFIPLSCISASETDTLKLPNLTFGGREIEFIVTP
                     PSSIAAQYITYANVSNYRGNFTISSCTQDQWFSRGLSTTNSSGAFVESMTSFTAIDNA
                     DLFFCNNYCTHQGGGGAINATGLISFKNNQNILFYNNTTIGTQFTGVALRTERNRGGA
                     LYGSSIELINNHSLNFINNTSGDMGGAVSTIQNLVIKNTSGIVAFENNHTTDHIPNTF
                     ATILARGGAVGCQGACEISHNTGPVVFNSNYGGYGGAISTGGQCIFRDNKDKLIFINN
                     SALGWHNTSAQGNGAVISAGGEFGLLNNKGPIYFENNNASYIAGAISCNNLNFQENGP
                     IYFLNNSALYGGAFHLFASPAANYIHTGSGDIIFNNNTELSTTGMSAGLRKLFYIPGT
                     TNNNPITLSLGAKKDTRIYFYDLFQWGGLKKANTPPENSPHTVTINPSDEFSGAVVFS
                     YKNISSDLQAHMIASKTHNQIKDSPTTLKFGTMSIENGAEFEFFNGPLTQESTSLLAL
                     GQDSILTVGKDASLTITHLGIILPGLLNDQGTTAPRIRVNPQDMTQNTNSNQAPVSTE
                     NVATQKIFFSGLVSLVDENYESVYDSCDLSRGKANQPILHIETTNDAQLSNDWKNTLN
                     TSLYSLPHYGYQGLWTSNWMTTTRTVSLTNSTETQTANNSIQEQKNTSETFDSNSTTT
                     AKIPSIRASTGGTTPLATTDVTVTRHSLVVSWTPIGYIADPARRGDLIANNLVSSGRN
                     TTLYLRSLLPDDSWFALQGSAATLFTKQQKRLDYHGYSSASKGYAISSQASGAHGHKF
                     LFSFSQSSDTMKEKRTNNKISSRYYLSALCFEQPMFDRIALIGAAAYNYGTHKTYNFY
                     GTKKFSKGNFHSTTLGGSLRCELRDSMPFQSIMLTPFIQALISRTEPASIQEQGDLAR
                     LFSLKQPHTAVVSPIGIKGVYSSNKWPTVSCEMEVAYQPTLYWKRPILNTVLIKNNGS
                     WETTNTPLAKHSFYGRGSSSLKFSYLKLFANYQAQVATSTVSHYMNAGGALVF"
gene                 8307. .11270
                     /gene="TC0263"
CDS                  8307. .11270
                     /gene="TC0263"
                     /note="similar to GP:4376736; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="polymorphic membrane protein G family"
                     /protein_id="AAF39132.1"
                     /db_xref="GI:7190303"
                     /translation="MMQTPFHKFFLLAMLSYSLLQGGHAADISMPPGIYDGTTLTAPF
                     PYTVIGDPRGTKVTSSGSLELKNLDNSIATLPLSCFGNLLGNFTIAGRGHSLVFENIR
                     TSTNGAALSNHAPSGLFVIEAFDELSLLNCNSLVSVVPQTGGTTTSVPSNGTIYSRTD
                     LVLRDIKKVSFYSNLVSGDGGAIDAQSLMVNGIEKLCTFQENVAQSDGGACQVTKTFS
                     AVGNKVPLSFLGNVAGNKGGGVAAVKDGQGAGGATDLSVNFANNTAVEFEGNSARIGG
                     GIYSDGNISFLGNAKTVFLSNVASPIYVDPAAAGGQPPADKDNYGDGGAIFCKNDTNI
                     GEVSFKDEGVVFFSKNIAAGKGGAIYAKKLTISDCGPVQFLGNVANDGGAIYLVDQGE
                     LSLSADRGDIIFDGNLKRMATQGAATVHDVMVASNAISMATGGQITTLRAKEGRRILF
                     NDPIEMANGQPVIQTLTVNEGEGYTGDIVFAKGDNVLYSSIELSQGRIILREQTKLLV
                     NSLTQTGGSVHMEGGSTLDFAVTTPPAANSMALTNVHFSLASLLKNNGVTNPPTNPPV
                     QVSSPAVIGNTAAGTVTISGPIFFEDLDETAYDNNQWLGADQTIDVLQLHLGANPPAN
                     APTDLTLGNESSKYGYQGSWTLQWEPDPANPPQNNSYMLKASWTKTGYNPGPERVASL
                     VSNSLWGSILDVRSAHSAIQASIDGRAYCRGIWISGISNFFYHDQDALGQGYRHISGG
                     YSIGANSYFGSSMFGLAFTETFGRSKDYVVCRSNDHTCVGSVYLSTRQALCGSCLFGD
                     AFVRASYGFGNQHMKTSYTFAEESNVRWDNNCVVGEVGAGLPIMLAASKLYLNELRPF
                     VQAEFAYAEHESFTERGDQAREFKSGHLMNLSIPVGVKFDRCSSKHPNKYSFMGAYIC
                     DAYRSISGTETTLLSHKETWTTDAFHLARHGVMVRGSMYASLTGNIEVYGHGKYEYRD
                     ASRGYGLSIGSKIRF"
gene                 11292. .14243
                     /gene="TC0264"
CDS                  11292. .14243
                     /gene="TC0264"
                     /note="similar to GP:4376737; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="polymorphic membrane protein H family"
                     /protein_id="AAF39133.1"
                     /db_xref="GI:7190304"
                     /translation="MLVMPFSLRSTSFCFLACLCSYSYGLASSPQVLTPNVIIPFKGD
                     DIYLNGDCVFASIYAGAEQGSIISANGQNLTIVGQNHTLSFTDSQGPALQNCAFISAE
                     EKISLRDFSSLLFSKNVSCGEKGMISGKTVSISGGDSIVFKDNSVGYSSLPSVGQTPT
                     TPIVGDVLKGSIFCVETGLEISGVKKELVFDNTAGNFGAVFCSRAAQGDTTFTVKDCK
                     GKILFQDNVGSCGGGVIYKGEVLFQDNEGEMLFRGNSAHDDLGILDANPQPPTEVGGG
                     GGVICTPEKTVTFKGNKGPITFDYNFAKGRGGAIQSQTFSLVADSAVVFSNNTAEKGG
                     GAIYALEVNVSTNGGSILFEGNRASEGGAICVSEPIAANNGGLTLHAADGDIIFSKNM
                     TSDRPGERSAIRILDSGTNVSLNASGASKMIFYDPVVQNNPATPPTGTSGEIKINESG
                     SGSVVFTAETLTPSEKLNVINATSNFPGNLTVSSGELVVTKGATLTVGNITATSGRVT
                     LGSGASLSAVAGTAGTCTVSKLGIDLESFLVPTYETAKLGADTTVAVNNNPTLDLVMA
                     NETEMYDNPLFMNAVTIPFVTLVSLQTTGGVTTSAVTLNNADTAHYGYQGSWSADWRR
                     PPLAPDPSGMTPLDKSNTLYVTWRPSSNYGVYKLDPQRRGELVPNSLWVSGSALRTFT
                     NGLKEHYVSRDVGFIASVQALGDYVLNYKQGNRDGFLARYGGFQAVAASHYENGGIFG
                     VAFGQLYGQTKSRLYDSKDAGNITILSCFGRSYIDVKGTETVVYWETAYGYSVHRMHT
                     QYFNGKTNKFDHSKCRWHNNSYYAFVGAEHNFLEYCIPTRQLARDYDLTGFMRFEMSG
                     GWSSGAKETGALPRHFDRGTGHNMSLPIGVVAHAVSNGRRSPPSKLTINMGYRPDIWR
                     VTPHCNMKIIANGVKTPIQGSPLARHAFFLEVHDTLYVRHLGRAYMNYSLDARHRQTT
                     HFVSLGLNRIF"
BASE COUNT     4185 a   2517 c   3329 g   4451 t
ORIGIN

  Query Match           55.2%;  Score 1599.4;  DB 1;  Length 14482;
  Best Local Similarity   72.6%;  Pred. No. 0;
  Matches 2128;  Conservative    0;  Mismatches  766;  Indels   39;  Gaps    3;

Qy      1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATTGGAAACTCCCTATCAGGACTAGCTAGAGAG 60
          |||||||||  ||| || ||| ||||||||||||| |||  ||   ||| |  || ||||
Db   5055 ATGAAAAAACTGTTCTTTTTTGTCCTTATTGGAAGCTCTATACTGGGATTTACTCGAGAA 4996

Qy     61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120
          || ||| ||   ||  || | | ||| |     |  |||| | | || ||  | ||   |
Db   4995 GTCCCTCCTTCGATTCTTTTAAAGCCTATACTAAATCCATACCATATGACCGGGTTATTT 4936

Qy    121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180
          |      ||  ||| ||||   |||||||| || |||||||| | | | || | |||||
Db   4935 TTTCCCAAGGTTAATTTGCTTGGAGACACACATAATCTCACTGATTACCATTTGGATAAT 4876

Qy    181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240
          |||   | ||| ||||||   ||||||| ||||||  |||||||||||||| ||||| ||
Db   4875 CTAAAATGCATTCTGGCTTGCCTACAAAGAACTCCTTATGAAGGAGCTGCTTTCACAGTA 4816

Qy    241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300
          || |||||| || | ||||  |||||||||| || |||||| || |   ||||||||
Db   4815 ACCGATTACTTAGGTTTTTCAGATACACAAAAGGATGGTATTTTTTGTTTTAAAAATCTT 4756

Qy    301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360
          || || || ||||| || |  |||||||   | | ||  ||  |||||||  |  | |||
Db   4755 ACTCCAGAGAGTGGAGGGGTTATTGGTTCCCCAACTCAAAACACTCCTACTATAAAAATT 4696

Qy    361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACCATTTACATCG 420
          | | ||||||| || || ||  | || ||||||||||| |||   ||||  |  |||  |
```


Db 4695 CATAATACAATCGGCCCCGTTCTTTTCGAAAATAATACCTGTCATAGACTGTGGACACAG 4636
Qy 421 AGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 480
Db 4635 ACCGATCCCGAAAATGAAGGAAACAAAGCACGCGAAGGCGGGGCAATTCATGCTGGGGAC 4576
Qy 481 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCGA 540
Db 4575 GTTTACATAAGCAATAACCAGAACCTTGTCGGATTCATAAAGAACTTTGCTTATGTTCAA 4516
Qy 541 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 600
Db 4515 GGTGGAGCTATTAGT---GCTAATACTTTTGCCTATAAAGAAAATAAATCGAGCTTTCTT 4459
Qy 601 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 660
Db 4458 TGCCTAAATAACTCTTGTATACAAACTAAGACGGGAGGGAAAGGTGGTGCTATTTACGTT 4399
Qy 661 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTTATCAATAACGCCTGT 720
Db 4398 AGTACGAGCTGCTCTTTCGAGAACAATAACAAGGATCTGCTTTTCATCCAAAACTCCGGC 4339
Qy 721 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 780
Db 4338 TGTGCAGGAGGAGCTATCTTCTCTCCAACCTGTTCTCTAATAGGAAACCAAGGAGATATT 4279
Qy 781 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 840
Db 4278 GTTTTTTACAGCAACCACGGTTTTAAAAATGTTGATAATGCAACTAACGAATCTGGGGAT 4219
Qy 841 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 900
Db 4218 GGAGGAGCTATTAAAGTAACTACCCGCTTGGACATCACCAATAATGGTAGTCAAATCTTT 4159
Qy 901 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 960
Db 4158 TTTTCTGATAATATCTCAAGAAATTTTGGAGGAGCTATTCATGCTCCTTGTCTTCATCTT 4099
Qy 961 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1020
Db 4098 GTTGGTAATGGGCCAACCTATTTTACAAACAATATAGCTAATCACACAGGTGGGGCTATT 4039
Qy 1021 TATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1080
Db 4038 TATATAACAGGAACAGAAACCTCAAAGATTTCTGCAGATCACCATGCTATTATTTTTGAT 3979
Qy 1081 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1140
Db 3978 AATAACATTTCTGCAAACGCCACCAATGCGGACGGATCTAGCAGCAACACTAATCCTCCT 3919
Qy 1141 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1200
Db 3918 CACAGAAATGCGATCACTATGGACAATTCCGCTGGAGGAATAGAACTTGGTGCAGGGAAG 3859
Qy 1201 AGCCAAAATTTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCC 1260
Db 3858 AGCCAGAATCTTATTTTCTATGATCCTATTCAAGTGACGAATGCTGGAGTTACCGTAGAC 3799
Qy 1261 TTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCT 1320
Db 3798 TTCAATAAGGATGCCTCCCAAACCGGATGTGTAGTTTTCTCTGGAGCGACTGTCCTTTCT 3739
Qy 1321 GCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAAT 1380
Db 3738 GCAGATATTTCTCAGGCTAATTTGCAAACTAAAACACCTGCAACGCTTACTCTCAGTCAC 3679
Qy 1381 GGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGG 1440
Db 3678 GGTCTTCTGTGTATCGAAGATCGTGCTCAGCTCACAGTGAACAATTTTACACAAACAGGA 3619
Qy 1441 GGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTA---------- 1481
Db 3618 GGGATTGTAGCCTTAGGAAATGGAGCAGTTTTAAGCAGCTACCAACACAGCACTACAGAC 3559
Qy 1482 -----TAAAAATGGTGCAGGAAATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATT 1536
Db 3558 GCCACTCAAACTCCCCCTACAACCACCACTACAGATGCTTCCGTAACTCTTAATCACATT 3499
Qy 1537 GGATTGAATCTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAG 1596
Db 3498 GGATTAAATCTCCCCTCTATTCTTAAGGATGGAGCAGAGATGCCTCTATTATGGGTAGAA 3439
Qy 1597 CCTACAA------------ATAACAGCAATAACTATACAGCAGATACTGCAGCTACCTTT 1644
Db 3438 CCTATAAGCACAACTCAAGGTAACACTACAACATATACGTCAGATACCGCGGCTTCCTTC 3379
Qy 1645 TCATTAAGTGATGTAAAACTCTCACTCATTGATGACTATGGGAATTCTCCTTATGAATCC 1704
Db 3378 TCATTAAATGGAGCCACACTCTCTCTCATTGATGAAGATGGAAATTCTCCCTATGAAAAC 3319
Qy 1705 ACAGATCTAACCCATGCTCTGTCATCACAGCCTATGCTATCTATTTCTGAGGCTAGTGAT 1764
Db 3318 ACGGACCTCTCTCGTGCATTGTACGCTCAACCTATGCTAGCAATTTCTGAGGCCAGTGAT 3259
Qy 1765 AACCAGCTAAGATCTGATGATATGGATTTTTCGGGACTAAATGTCCCTCATTATGGATGG 1824
Db 3258 AACCAATTGCAATCCGAAAGCATGGACTTTTCTAAAGTTAATGTTCCTCACTATGGATGG 3199
Qy 1825 CAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAAGATCCAGAACCAGCATCTTCAGCA 1884
Db 3198 CAAGGACTTTGGACCTGGGGGTGGGCAAAAACTGAAAATCCAACAACAACTCCTCCAGCA 3139
Qy 1885 ACAATCACAGATCCACAAAAAGCCAATAGATTCCATAGAACCTTATTACTGACTTGGCTT 1944
Db 3138 ACAATTACTGATCCGAAAAAAGCTAATCAGTTTCATAGAACTTTATTATTAACGTGGCTC 3079
Qy 1945 CCTGCTGGGTATGTTCCTAGCCCGAAACACAGAAGTCCCCTCATAGCGAATACCTTATGG 2004
Db 3078 CCTGCTGGTTATATCCCCAGCCCTAAACATAAAAGCCCTTTAATAGCTAATACCTTGTGG 3019
Qy 2005 GGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAATAGTGCAGAACTGACACCTAGTGAT 2064
Db 3018 GGGAATATACTTTTTGCAACGGAAAACTTAAAAAATAGCTCAGGGCAAGAACTTCTTGAT 2959
Qy 2065 CATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGA 2124
Db 2958 CGTCCTTTCTGGGGAATTACAGGAGGGGGCTTGGGGATGATGGTCTATCAAGAACCTAGA 2899
Qy 2125 GAAAATCATCCTGGATTCCATATGCGCTCTTCCGGATACTCTGCGGGGATGATAGCAGGG 2184
Db 2898 AAAGACCATCCTGGATTCCACATGCATACCTCCGGATATTCAGCAGGAATGATTACAGGA 2839
Qy 2185 CAGACACACACCTTCTCATTGAAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTAC 2244
Db 2838 AACACACATACCTTCTCATTACGATTCAGCCAGTCCTATACAAAACTCAATGAACGTTAT 2779
Qy 2245 GCAAAAAACAACGTATCTTCTAAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTG 2304
Db 2778 GCCAAGAACTATGTGTCTTCTAAAAATTACTCTTGCCAAGGGGAAATGCTTTTGTCCTTA 2719
Qy 2305 CAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGT 2364
Db 2718 CAAGAAGGACTCATGCTGACTAAACTAATTGGTCTCTATAGTTATGGGAATCACAACAGC 2659
Qy 2365 CACCATTTCTATACCCAAGGAGAAAATCTAACATCTCAAGGGACGTTCCGTAGTCAAACG 2424
Db 2658 CACCATTTCTATACCCAAGGAGAAGACCTATCGTCTCAAGGGGAGTTCCATAGTCAGACT 2599
Qy 2425 ATGGGAGGTGCTGTTTTTTTTGATCTCCCTATGAAACCCTTTGGATCAACGCATATACTG 2484
Db 2598 TTTGGAGGGGCTGTCTTTTTTGATCTACCTCTGAAACCTTTTGGAAGAACACACATACTT 2539
Qy 2485 ACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTG 2544
Db 2538 ACAGCTCCTTTCTTAGGTGCCATTGGTATGTATTCTAAGCTGTCTAGCTTTACAGAAGTA 2479

```
Qy   2545 GGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATT 2604
          ||||||||||| ||| |||| |||| | || ||||| |||||||||||  | |||||
Db   2478 GGAGCCTATCCAAGAACCTTTATTACAGAAACGCCTTTAATCAATGTCCTGATTCCTATC 2419

Qy   2605 GGAGTTAAAGGTAGCTTTATGAATGCTACCCAAAGACCTCAAGCCTGGACTGTAGAATTG 2664
          ||||| |||||||||||| |||||||| ||||| |||||||| ||||||||||||||  |
Db   2418 GGAGTAAAAGGTAGCTTCATGAATGCCACCCATAGACCTCAGGCCTGGACTGTAGAGCTT 2359

Qy   2665 GCATACCAACCCGTTCTGTATAGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAGT 2724
          || |||||||| ||||| || |||||||||||  | ||| | |||||  | || ||  ||
Db   2358 GCTTACCAACCTGTTCTTTACAGACAAGAACCTAGTATCTCTACCCAATTACTCGCTGGT 2299

Qy   2725 AAGGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATC 2784
          || ||||| ||||||||   |||||| ||  |||| || || ||  |  | ||||||||
Db   2298 AAAGGTATGTGGTTTGGGCATGGAAGTCCTGCATCTCGCCACGCTCTAGCTTATAAAATT 2239

Qy   2785 TCACAGCAAACACAACCTTTGAGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTAC 2844
          |||||| ||||||| | |||| | |    ||| || ||  |||||||||| |||| |||
Db   2238 TCACAGAAAACACAGCTTTTGCGATTTGCAACACTTCAACTCCAGTATCACGGATACTAT 2179

Qy   2845 TCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAAATTGCTCTGCGATTCTA 2897
          || ||||| || |||||||||||||| ||||| ||  |  || | || |||||
Db   2178 TCGTCTTCCACTTTCTGTAATTATCTGAATGGAGAGGTATCTTTACGTTTCTA 2126
```

RESULT 7
AX338273/c
LOCUS AX338273 1634 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 43 from Patent WO0181379.
ACCESSION AX338273
VERSION AX338273.1 GI:18128808
KEYWORDS .
SOURCE Chlamydia trachomatis.
ORGANISM Chlamydia trachomatis
Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE 1
AUTHORS Bhatia,A., Probst,P. and Stromberg,E.J.
TITLE Compounds and methods for treatment and diagnosis of chlamydial infection
JOURNAL Patent: WO 0181379-A 43 01-NOV-2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1. .1634
/organism="Chlamydia trachomatis"
/db_xref="taxon:813"
BASE COUNT 485 a 293 c 356 g 500 t
ORIGIN

Query Match 41.1%; Score 1190.6; DB 6; Length 1634;
Best Local Similarity 99.0%; Pred. No. 3.4e-264;
Matches 1209; Conservative 0; Mismatches 9; Indels 3; Gaps 1;

```
Qy      1 ATGAAAAAAGCGTTTTTCTTTTTCCTTATTGGAAACTCCCTATCAGGACTAGCTAGAGAG 60
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   1218 ATGAAAAAAGCGTTTTTCTTTTTCCTTATCGGAAACTCCCTATCAGGACTAGCTAGAGAG 1159

Qy     61 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1158 GTTCCTTCTAGAATCTTTCTTATGCCCAACTCAGTTCCAGATCCTACGAAAGAGTCGCTA 1099

Qy    121 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1098 TCAAATAAAATTAGTTTGACAGGAGACACTCACAATCTCACTAACTGCTATCTCGATAAC 1039

Qy    181 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1038 CTACGCTACATACTGGCTATTCTACAAAAAACTCCCAATGAAGGAGCTGCTGTCACAATA 979

Qy    241 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    978 ACAGATTACCTAAGCTTTTTTGATACACAAAAAGAAGGTATTTATTTTGCAAAAAATCTC 919

Qy    301 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    918 ACCCCTGAAAGTGGTGGTGCGATTGGTTATGCGAGTCCCAATTCTCCTACCGTGGAGATT 859

Qy    361 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACCATTTACATCG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    858 CGTGATACAATAGGTCCTGTAATCTTTGAAAATAATACTTGTTGCAGACTATTTACATGG 799

Qy    421 AGTAATCCTAATGCAGCTGTTAATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 480
          || |||||| |||| ||||   ||||||||||||||||||||||||||||||||||||||
Db    798 AGAAATCCTTATGCTGCTG---ATAAAATAAGAGAAGGCGGAGCCATTCATGCTCAAAAT 742

Qy    481 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCGA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    741 CTTTACATAAATCATAATCATGATGTGGTCGGATTTATGAAGAACTTTTCTTATGTCCAA 682

Qy    541 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    681 GGAGGAGCCATTAGTACCGCTAATACCTTTGTTGTGAGCGAGAATCAGTCTTGTTTTCTC 622

Qy    601 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    621 TTTATGGACAACATCTGTATTCAAACTAATACAGCAGGAAAAGGTGGCGCTATCTATGCT 562

Qy    661 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTTATCAATAACGCCTGT 720
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    561 GGAACGAGCAATTCTTTTGAGAGTAATAACTGCGATCTCTTCTTCATCAATAACGCCTGT 502

Qy    721 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    501 TGTGCAGGAGGAGCGATCTTCTCCCCTATCTGTTCTCTAACAGGAAATCGTGGTAACATC 442

Qy    781 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    441 GTTTTCTATAACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGAT 382

Qy    841 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    381 GGAGGAGCAATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTT 322

Qy    901 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    321 TTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTA 262

Qy    961 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    261 GTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATC 202

Qy   1021 TATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 1080
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    201 TATATAGACGGAACCAGTAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAAT 142

Qy   1081 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     141 GAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCT 82

Qy   1141 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      81 AGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGT 22

Qy   1201 AGCCAAAATTTAATTTTTTAT 1221
          ||||||||||||||||||||
Db      21 AGCCAAAATTTAATTTTTTAT 1
```

RESULT 8
AX349485
LOCUS AX349485 2805 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 8 from Patent WO0202606.
ACCESSION AX349485
VERSION AX349485.1 GI:18615349
KEYWORDS .
SOURCE Chlamydophila pneumoniae.
ORGANISM Chlamydophila pneumoniae

```
            Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
REFERENCE   1
  AUTHORS   Ratti,G. and Grandi,G.
  TITLE     Immunisation against Chlamydia pneumoniae
  JOURNAL   Patent: WO 0202606-A 8 10-JAN-2002;
            Chiron S.p.A. (IT)
FEATURES             Location/Qualifiers
     source          1. .2805
                     /organism="Chlamydophila pneumoniae"
                     /db_xref="taxon:83558"
BASE COUNT      812 a    696 c    527 g    770 t
ORIGIN

  Query Match             4.7%;  Score 136.4;  DB 6;  Length 2805;
  Best Local Similarity   45.9%;  Pred. No. 8.2e-21;
  Matches  762;  Conservative    0;  Mismatches  856;  Indels   42;  Gaps    7;

Qy   1253 CTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTG 1312
Db   1172 CTCTAATCTTTAATCCCAATGCGAACCATCAGGGAACGATCTTATTTTCTTCAGCCTATA 1231

Qy   1313 TTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTC 1372
Db   1232 TCCCAGAAGCTTCTGACTACGAAAATAATTTCATTAGCAGCTCGAAAAATACCTCTGAAC 1291

Qy   1373 TCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACAC 1432
Db   1292 TTCGCAATGGTGTCCTCTCTATCGAGGATCGTGCGGGATGGCAATTCTATAAGTTCACTC 1351

Qy   1433 AAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTG 1492
Db   1352 AAAAAGGAGGTATCCTTAAATTAGGGCATGCGGCGAGTATTGCAACAACTGCCAACTCTG 1411

Qy   1493 CAGGAAATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTT 1552
Db   1412 AGACTCCATCAACTAGTGTAGGCTCCCAGGTCATCATTAATAACCTTGCGATTAACCTCC 1471

Qy   1553 CCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCA 1612
Db   1472 CCTCGATCTTAGCAAAAGGAAAAGCTCCTACCTTGTGGATCCGTCCTCTACAATCTAGTG 1531

Qy   1613 ATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCA 1672
Db   1532 CTCCTTTCACAGAGGACAATAACCCTACAATTACTTTATCAGGTCCT---CTGACACTCT 1588

Qy   1673 TTGATGACTATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTGTCATCAC 1732
Db   1589 TAAATGAGGAAAACCGCGATCCCTACGACAGTATAGATCTCTCTGAGCCTTTACAAAACA 1648

Qy   1733 AGCCTATGCTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGATATGGATT 1792
Db   1649 TTCATCTTCTTTCTTTATCGGATGTAACAGCACGTCATATCAATACCGATAACTTTCATC 1708

Qy   1793 TTTCGGGACTAAATGTCCCT---CATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGG 1849
Db   1709 CTGAAAGCTTAAATGCGACTGAGCATTACGGTTATCAAGGCATCTGGTCTCCTTATTGGG 1768

Qy   1850 CAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAAGCCA 1909
Db   1769 --------------TAGAGACGATAACAACAACAAATAACGCTTCTATAGAGACGGCAA 1813

Qy   1910 ATAGATTCCATAGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGCCCGA 1969
Db   1814 ACACCCTCTACAGAGCTCTGTATGCCAATTGGACTCCCTTAGGATATAAGGTCAATCCTG 1873

Qy   1970 AACACAGAAGTCCCCTCATAGCGAATACCTTATGG------GGGAATATGCTGCTTGCAA 2023
Db   1874 AATACCAAGGAGATCTTGCTACGACTCCCCTATGGCAATCCTTTCATACTATGTTCTCTC 1933

Qy   2024 CAGAAAGCTTAAAAAATAGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTA 2083
Db   1934 TATTAAGAAGTTATAATCGAACTGGTGATTCTGATATCGAGAGGCCTTTCTTAGAAATTC 1993

Qy   2084 CAGGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCC 2143
Db   1994 AAGGGATTGCCGACGGCCTCTTTGTTCATCAAAATAGCATCCCCGGGGCTCCAGGATTCC 2053

Qy   2144 ATATGCGCTCTTCCGGATACTC------TGCGGGGATGATAGCAGGGCAGACACACACCT 2197
Db   2054 GTATCCAATCTACAGGGTATTCCTTACAAGCATCCTCCGAAACTTCTTTACATCAGAAAA 2113

Qy   2198 TCTCATTGAAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTACGCAAAAAACAACG 2257
Db   2114 TCTCCTTAGGTTTTGCACAGTTCTTCACCCGCACTAAAGAAATCGGATCAAGCAACAACG 2173

Qy   2258 TATCTTCTAAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCT 2317
Db   2174 TCTCGGCTCACAATACAGTCTCTTCACTTTATGTTGAGCTTCCGTGGTTCCAAGAGGCCT 2233

Qy   2318 TGCTGACTAAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCACCATTTCTATA 2377
Db   2234 TTGCAACATCCACAGTGTTAGCGTATGGCTATGGGGACCATCACCTCCACAGCCTACATC 2293

Qy   2378 CCCAAGGAGAAAATCTAACATCTCAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTG 2437
Db   2294 CCTCACATCAAGAACAGGCA---GAAGGGACGTGTTATAGCCATACATTAGCAGCAGCTA 2350

Qy   2438 TTTTTTTTGATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCTCCCTTTT 2497
Db   2351 TCGGCTGTTCTTTCCCTTGGCAAC------AGAAATCCTATCTTCACCTCAGCCCGTTCG 2404

Qy   2498 TAGGTGCTCTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGC 2557
Db   2405 TTCAGGCAATTGCAATACGTTCTCACCAAACAGCGTTCGAAGAGATTGGTGACAATCCCC 2464

Qy   2558 GAAGCTTTTCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTA 2617
Db   2465 GAAAGTTTGTCTCTCAAAAGCCTTTCTATAATCTGACCTTACCTCTAGGAATCCAAGGAA 2524

Qy   2618 GCTTTATGAATGCTACCCAAAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCG 2677
Db   2525 AATGGCAGTCAAAATTCCACGTACCTACAGAATGGACTCTAGAACTTTCTTACCAACCGG 2584

Qy   2678 TTCTGTATAGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGT 2737
Db   2585 TACTCTATCAACAAAATCCCCAAATCGGTGTCACGCTACTTGCGAGCGGAGGTTCCTGGG 2644

Qy   2738 TTGGTAGTGGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAGCAAACAC 2797
Db   2645 ATATCCTAGGCCATAACTATGTTCGCAATGCTTTAGGGTACAAAGTCCACAATCAAACTG 2704

Qy   2798 AACCTTTGAGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCTTCAACCT 2857
Db   2705 CGCTCTTCCGTTCTCTCGATCTATTCTTGGATTACCAAGGATCGGTCTCCTCCTCGACAT 2764

Qy   2858 TCTGTAATTATCTCAATGGGGAAATTGCTCTGCGATTCTA 2897
Db   2765 CTACGCACCATCTCCAAGCAGGAAGTACCTTAAAATTCTA 2804


RESULT 9
AE002190/c
LOCUS       AE002190               10746 bp    DNA     linear   BCT 30-MAY-2000
DEFINITION  Chlamydophila pneumoniae AR39, section 23 of 94 of the complete
            genome.
ACCESSION   AE002190 AE002161
VERSION     AE002190.2  GI:8163403
KEYWORDS    .
SOURCE      Chlamydophila pneumoniae AR39.
  ORGANISM  Chlamydophila pneumoniae AR39
            Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
REFERENCE   1  (bases 1 to 10746)
  AUTHORS   Read,T.D., Brunham,R.C., Shen,C., Gill,S.R., Heidelberg,J.F.,
            White,O., Hickey,E.K., Peterson,J., Umayam,L.A., Utterback,T.,
            Berry,K., Bass,S., Linher,K., Weidman,J., Khouri,H., Craven,B.,
            Bowman,C., Dodson,R., Gwinn,M., Nelson,W., DeBoy,R., Kolonay,J.,
```

```
            McClarty,G., Salzberg,S.L., Eisen,J. and Fraser,C.M.
  TITLE     Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
            pneumoniae AR39
  JOURNAL   Nucleic Acids Res. 28 (6), 1397-1406 (2000)
  MEDLINE   20150255
   PUBMED   10684935
REFERENCE   2  (bases 1 to 10746)
  AUTHORS   Read,T.D., Brunham,R.C., Shen,C., Gill,S.R., Heidelberg,J.F.,
            White,O., Hickey,E.K., Peterson,J., Umayam,L.A., Utterback,T.,
            Berry,K., Bass,S., Linher,K., Weidman,J., Khouri,H., Craven,B.,
            Bowman,C., Dodson,R., Gwinn,M., Nelson,W., DeBoy,R., Kolonay,J.,
            McClarty,G., Salzberg,S.L., Eisen,J. and Fraser,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAR-2000) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
COMMENT     On Jun 1, 2000 this sequence version replaced gi:7189209.
FEATURES             Location/Qualifiers
     source          1. .10746
                     /organism="Chlamydophila pneumoniae AR39"
                     /strain="AR39"
                     /db_xref="taxon:115711"
                     /note="synonym: Chlamydia pneumoniae AR39"
     gene            complement(115. .2973)
                     /gene="CP0285"
     CDS             complement(115. .2973)
                     /gene="CP0285"
                     /note="similar to GP:4376752; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="polymorphic membrane protein E/F family"
                     /protein_id="AAF38142.1"
                     /db_xref="GI:7189210"
                     /translation="MSKTPPKFLFYLGNFTACMFGMTPAVYSLQTDSLEKFALERDEE
                     FRTSFPLLDSLSTLTGFSPITTFVGNRHNSSQDIVLSNYKSIDNILLLWTSAGGAVSC
                     NNFLLSNVEDHAFFSKNLAIGTGGAIACQGACTITKNRGPLIFFSNRGLNNASTGGET
                     RGGAIACNGDFTISQNQGTFYFVNNSVNNWGGALSTNGHCRIQSNRAPLLFFNNTAPS
                     GGGALRSENTTISDNTRPIYFKNNCGNNGGAIQTSVTVAIKNNSGSVIFNNNTALSGS
                     INSGNGSGGAIYTTNLSIDDNPGTILFNNNYCIRDGGAICTQFLTIKNSGHVYFTNNQ
                     GNWGGALMLLQDSTCLLFAEQGNIAFQNNEVFLTTFGRYNAIHCTPNSNLQLGANKGY
                     TTAFFDPIEHQHPTTNPLIFNPNANHQGTILFSSAYIPEASDYENNFISSSKNTSELR
                     NGVLSIEDRAGWQFYKFTQKGGILKLGHAASIATTANSETPSTSVGSQVIINNLAINL
                     PSILAKGKAPTLWIRPLQSSAPFTEDNNPTITLSGPLTLLNEENRDPYDSIDLSEPLQ
                     NIHLLSLSDVTARHINTDNFHPESLNATEHYGYQGIWSPYWVETITTTNNASIETANT
                     LYRALYANWTPLGYKVNPEYQGDLATTPLWQSFHTMFSLLRSYNRTGDSDIERPFLEI
                     QGIADGLFVHQNSIPGAPGFRIQSTGYSLQASSETSLHQKISLGFAQFFTRTKEIGSS
                     NNVSAHNTVSSLYVELPWFQEAFATSTVLAYGYGDHHLHSLHPSHQEQAEGTCYSHTL
                     AAAIGCSFPWQQKSYLHLSPFVQAIAIRSHQTAFEEIGDNPRKFVSQKPFYNLTLPLG
                     IQGKWQSKFHVPTEWTLELSYQPVLYQQNPQIGVTLLASGGSWDILGHNYVRNALGYK
                     VHNQTALFRSLDLFLDYQGSVSSSTSTHHLQAGSTLKF"
     gene            complement(3012. .5828)
                     /gene="CP0286"
     CDS             complement(3012. .5828)
                     /gene="CP0286"
                     /note="similar to GP:4376751; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="polymorphic membrane protein E/F family"
                     /protein_id="AAF38143.1"
                     /db_xref="GI:7189211"
                     /translation="MRFFCFGMLLPFTFVLANEGLQLPLETYITLSPEYQAAPQVGFT
                     HNQNQDLAIVGNHNDFILDYKYYRSNGGALTCKNLLISENIGNVFFEKNVCPNSGGAI
                     YAAQNCTISKNQNYAFTTNLVSDNPTATAGSLLGGALFAINCSITNNLGQGTFVDNLA
                     LNKGGALYTETNLSIKDNKGPIIIKQNRALNSDSLGGGIYSGNSLNIEGNSGAIQITS
                     NSSGSGGGIFSTQTLTISSNKKLIEISENSAFANNYGSNFNPGGGGLTTTFCTILNNR
                     EGVLFNNNQSQSNGGAIHAKSIIIKENGPVYFLNNTATRGGALLNLSAGSGNGSFILS
                     ADNGDIIFNNNTASKHALNPPYRNAIHSTPNMNLQIGARPGYRVLFYDPIEHELPSSF
                     PILFNFETGHTGTVLFSGEHVHQNFTDEMNFFSYLRNTSELRQGVLAVEDGAGLACYK
                     FFQRGGTLLLGQGAVITTAGTIPTPSSTPTTVGSTITLNHIAIDLPSILSFQAQAPKI
                     WIYPTKTGSTYTEDSNPTITISGTLTLRNSNNEDPYDSLDLSHSLEKVPLLYIVDVAA
                     QKINSSQLDLSTLNSGEHYGYQGIWSTYWVETTTITNPTSLLGANTKHKLLYANWSPL
                     GYRPHPERRGEFITNALWQSAYTALAGLHSLSSWDEEKGHAASLQGIGLLVHQKDKNG
                     FKGFRSHMTGYSATTEATSSQSPNFSLGFAQFFSKAKEHESQNSTSSHHYFSGMCIEN
                     TLFKEWIRLSVSLAYMFTSEHTHTMYQGLLEGNSQGSFHNHTLAGALSCVFLPQPHGE
                     SLQIYPFITALAIRGNLAAFQESGDHAREFSLHRPLTDVSLPVGIRASWKNHHRVPLV
                     WLTEISYRSTLYRQDPELHSKLLISQGTWTTQATPVTYNALGIKVKNTMQVFPKVTLS
                     LDYSADISSSTLSHYLNVASRMRF"
     gene            6340. .6684
                     /gene="CP0287"
     CDS             6340. .6684
                     /gene="CP0287"
                     /note="hypothetical protein; identified by Glimmer2;
                     putative"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAF38144.1"
                     /db_xref="GI:7189212"
                     /translation="MSYYFSLWYLKVQQHFQAAFDFTRSLCSRISNFALGVIALLPII
                     GQLYVGLDWLLSRIKKPEFPSDVDQIVRVEHVVGHDHRSRVEDILKRQRLSLEPRDEG
                     KVRGDLPSAPFF"
     gene            6710. .7189
                     /gene="CP0288"
     CDS             6710. .7189
                     /gene="CP0288"
                     /note="hypothetical protein; identified by Glimmer2;
                     putative"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAF38145.1"
                     /db_xref="GI:7189213"
                     /translation="MFSEGTALNLFRIFAPLRNRVTTEYSRARQPDLHRIAIVYIGVL
                     DSESSKILERLISYMSCIYSESQMYLRFFMGKNVNQSAVLSKLHVENLHIRCGFFSED
                     AVPESEPFDLSIYVHTDRSCPLPTKKRSSSWELQTVELPESIYPQSEFLLMRPRMLS"
     gene            7195. .8364
                     /gene="CP0289"
     CDS             7195. .8364
                     /gene="CP0289"
                     /note="hypothetical protein; identified by Glimmer2;
                     putative"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAF38146.1"
                     /db_xref="GI:7189214"
                     /translation="MMKQGVGQDAKELYTFLSRGNEHYQPCLWFSLEEELGFLFDEKM
                     LCAPLSEDHYCHSYLVDLVDQHLKDLILSMFLDPQNISAGELLKVSINVGDSFSPLQQ
                     KDFLSMVLRDETGKNVVVVFKGVLSLPATQVCKLVEELNSKDYSYLNIFSCHGDSSPQ
                     LLFRKELEGTSGRYFTVICALYLGDTDMRSLQLASERIMVSREFDLVDAYAARCKLLK
                     IDHTNWRPGTFSRHADFADAVDVSAGFNSREFKLITQANQGILESGELPLPSKTFWEG
                     FLAFCDRVTVTRHFIPMLDAAIKQAVWTHKHPSLIDKECEALDLKTQCLPSIVSYLEY
                     VTNSHEKTSKGPFIQKEIIADCSPLKEALFPGSDEDVPSTSEDPSDDHPSDLEDS"
     gene            8500. .10518
                     /gene="CP0290"
     CDS             8500. .10518
                     /gene="CP0290"
                     /note="hypothetical protein; identified by Glimmer2;
                     putative"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAF38147.1"
                     /db_xref="GI:7189215"
                     /translation="MACYISIWISTVKQHFIRAFDFTRPLGSRITNFALGVIKAIPIL
                     GCVVIGVSWLVSTCSARRFGKPAFTSDVASIVKIEKTRGYNPLAWVEQYLRQLRVRLP
                     EGDLGKIHGKVSRDYVCDRTPQENLNMVPHQYLGELGRAFYGIRNRVTKAYQRVTPLE
                     VPCLTLVGFDILDPEDQVNFVRLANGIQTQYPQTQIKLYLISIQKIWNQCDGTISQEK
                     EQQLRSLGLDAKIKCVSAPALLLQKYLQSENLPSCDLLINYYGKQQSVRDVDSIKSLL
                     NLSSEHIPAISVTYRPDDPFYSYFFPGSQGGTAPDQRIPWSEQEHLQTYTTLSNPRC
                     DRYAVHLGMEDFASGVFLDPLRVSAPLSGEYSCPSYLLDLKSEELRCFLLSAFIDPNN
                     SGQGNPRPMSINFGNSPLGQRWSEFLSRVLHDETEKHVAVVCNNPQLIKKSFPSHSLS
                     LLENELEESGYSYLNIVSVSQERTCVKERRILSSDPSGRSFTVILTDLPEGSSDIRNL
                     QLASDRILVSSALDAADACASECKILEYEDPEQEWAQQYASFYRNIDRAGDLQRQGIP
                     GEPLGVSASTRVVLEKDIVFNLNAVIQQAMWKFKKRDLFAVESQALGDDMRRALEGYI
```

```
                    GSSLLVEGTIQPQVACNVNVSFATLDEAVCAACDSAQDAPSEENNTDD"
BASE COUNT      2938 a    2080 c    2490 g    3238 t
ORIGIN

  Query Match              4.7%;  Score 136.4;  DB 1;  Length 10746;
  Best Local Similarity   45.9%;  Pred. No. 6.9e-21;
  Matches  762;  Conservative    0;  Mismatches  856;  Indels   42;  Gaps    7;

Qy       1253 CTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTG 1312
Db       1748 CTCTAATCTTTAATCCCAATGCGAACCATCAGGGAACGATCTTATTTTCTTCAGCCTATA 1689

Qy       1313 TTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTC 1372
Db       1688 TCCCAGAAGCTTCTGACTACGAAAATAATTTCATTAGCAGCTCGAAAAATACCTCTGAAC 1629

Qy       1373 TCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACAC 1432
Db       1628 TTCGCAATGGTGTCCTCTCTATCGAGGATCGTGCGGGATGGCAATTCTATAAGTTCACTC 1569

Qy       1433 AAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTG 1492
Db       1568 AAAAAGGAGGTATCCTTAAATTAGGGCATGCGGCGAGTATTGCAACAACTGCCAACTCTG 1509

Qy       1493 CAGGAAATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTT 1552
Db       1508 AGACTCCATCAACTAGTGTAGGCTCCCAGGTCATCATTAATAACCTTGCGATTAACCTCC 1449

Qy       1553 CCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCA 1612
Db       1448 CCTCGATCTTAGCAAAAGGAAAAGCTCCTACCTTGTGGATCCGTCCTCTACAATCTAGTG 1389

Qy       1613 ATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCA 1672
Db       1388 CTCCTTTCACAGAGGACAATAACCCTACAATTACTTTATCAGGTCCT---CTGACACTCT 1332

Qy       1673 TTGATGACTATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTGTCATCAC 1732
Db       1331 TAAATGAGGAAAACCGCGATCCCTACGACAGTATAGATCTCTCTGAGCCTTTACAAAACA 1272

Qy       1733 AGCCTATGCTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGATATGGATT 1792
Db       1271 TTCATCTTCTTTCTTTATCGGATGTAACAGCACGTCATATCAATACCGATAACTTTCATC 1212

Qy       1793 TTTCGGGACTAAATGTCCCT---CATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGG 1849
Db       1211 CTGAAAGCTTAAATGCGACTGAGCATTACGGTTATCAAGGCATCTGGTCTCCTTATTGGG 1152

Qy       1850 CAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAAGCCA 1909
Db       1151 ---------------TAGAGACGATAACAACAACAAATAACGCTTCTATAGAGACGGCAA 1107

Qy       1910 ATAGATTCCATAGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGCCCGA 1969
Db       1106 ACACCCTCTACAGAGCTCTGTATGCCAATTGGACTCCCTTAGGATATAAGGTCAATCCTG 1047

Qy       1970 AACACAGAAGTCCCCTCATAGCGAATACCTTATGG------GGGAATATGCTGCTTGCAA 2023
Db       1046 AATACCAAGGAGATCTTGCTACGACTCCCCTATGGCAATCCTTTCATACTATGTTCTCTC 987

Qy       2024 CAGAAAGCTTAAAAAATAGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTA 2083
Db        986 TATTAAGAAGTTATAATCGAACTGGTGATTCTGATATCGAGAGGCCTTTCTTAGAAATTC 927

Qy       2084 CAGGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCC 2143
Db        926 AAGGGATTGCCGACGGCCTCTTTGTTCATCAAAATAGCATCCCCGGGGCTCCAGGATTCC 867

Qy       2144 ATATGCGCTCTTCCGGATACTC------TGCGGGGATGATAGCAGGGCAGACACACACCT 2197
Db        866 GTATCCAATCTACAGGGTATTCCTTACAAGCATCCTCCGAAACTTCTTTACATCAGAAAA 807

Qy       2198 TCTCATTGAAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTACGCAAAAAACAACG 2257
Db        806 TCTCCTTAGGTTTGCACAGTTCTTCACCCGCACTAAAGAAATCGGATCAAGCAACAACG 747

Qy       2258 TATCTTCTAAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCT 2317
Db        746 TCTCGGCTCACAATACAGTCTCTTCACTTTATGTTGAGCTTCCGTGGTTCCAAGAGGCCT 687

Qy       2318 TGCTGACTAAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCACCATTTCTATA 2377
Db        686 TTGCAACATCCACAGTGTTAGCGTATGGCTATGGGGACCATCACCTCCACAGCCTACATC 627

Qy       2378 CCCAAGGAGAAAATCTAACATCTCAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTG 2437
Db        626 CCTCACATCAAGAACAGGCA---GAAGGGACGTGTTATAGCCATACATTAGCAGCAGCTA 570

Qy       2438 TTTTTTTGATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCTCCCTTTT 2497
Db        569 TCGGCTGTTCTTTCCCTTGGCAAC------AGAAATCCTATCTTCACCTCAGCCCGTTCG 516

Qy       2498 TAGGTGCTCTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGC 2557
Db        515 TTCAGGCAATTGCAATACGTTCTCACCAAACAGCGTTCGAAGAGATTGGTGACAATCCCC 456

Qy       2558 GAAGCTTTTCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTA 2617
Db        455 GAAAGTTTGTCTCTCAAAAGCCTTTCTATAATCTGACCTTACCTCTAGGAATCCAAGGAA 396

Qy       2618 GCTTTATGAATGCTACCCAAAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCG 2677
Db        395 AATGGCAGTCAAAATTCCACGTACCTACAGAATGGACTCTAGAACTTTCTTACCAACCGG 336

Qy       2678 TTCTGTATAGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGT 2737
Db        335 TACTCTATCAACAAAATCCCCAAATCGGTGTCACGCTACTTGCGAGCGGAGGTTCCTGGG 276

Qy       2738 TTGGTAGTGGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAGCAAACAC 2797
Db        275 ATATCCTAGGCCATAACTATGTTCGCAATGCTTTAGGGTACAAAGTCCACAATCAAACTG 216

Qy       2798 AACCTTTGAGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCTTCAACCT 2857
Db        215 CGCTCTTCCGTTCTCTCGATCTATTCTTGGATTACCAAGGATCGGTCTCCTCCTCGACAT 156

Qy       2858 TCTGTAATTATCTCAATGGGGAAATTGCTCTGCGATTCTA 2897
Db        155 CTACGCACCATCTCCAAGCAGGAAGTACCTTAAAATTCTA 116


RESULT 10
AE001631
LOCUS       AE001631             12083 bp    DNA     linear   BCT 01-DEC-2000
DEFINITION  Chlamydia pneumoniae section 47 of 103 of the complete genome.
ACCESSION   AE001631 AE001363
VERSION     AE001631.1  GI:4376750
KEYWORDS    .
SOURCE      Chlamydophila pneumoniae CWL029.
  ORGANISM  Chlamydophila pneumoniae CWL029
            Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
REFERENCE   1  (bases 1 to 12083)
  AUTHORS   Kalman,S., Mitchell,W., Marathe,R., Lammel,C., Fan,J., Hyman,R.W.,
            Olinger,L., Grimwood,J., Davis,R.W. and Stephens,R.S.
  TITLE     Comparative genomes of Chlamydia pneumoniae and C. trachomatis
  JOURNAL   Nat. Genet. 21 (4), 385-389 (1999)
  MEDLINE   99206606
   PUBMED   10192388
REFERENCE   2  (bases 1 to 12083)
  AUTHORS   Kalman,S., Mitchell,W., Marathe,R., Lammel,C., Fan,J., Olinger,L.,
            Grimwood,J., Davis,R.W. and Stephens,R.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-DEC-1998) Program in Infectious Diseases, University
            of California, 235 Earl Warren Hall, Berkeley, CA 94720, USA
FEATURES             Location/Qualifiers
     source          1..12083
```

```
                     /organism="Chlamydophila pneumoniae CWL029"
                     /strain="CWL029"
                     /db_xref="taxon:115713"
     gene            287. .3103
                     /gene="pmp_15"
                     /note="CPn0466"
     CDS             287. .3103
                     /gene="pmp_15"
                     /codon_start=1
                     /transl_table=11
                     /product="Polymorphic Outer Membrane Protein E Family"
                     /protein_id="AAD18608.1"
                     /db_xref="GI:4376751"
                     /translation="MRFFCFGMLLPFTFVLANEGLQLPLETYITLSPEYQAAPQVGFT
                     HNQNQDLAIVGNHNDFILDYKYYRSNGGALTCKNLLISENIGNVFFEKNVCPNSGGAI
                     YAAQNCTISKNQNYAFTTNLVSDNPTATAGSLLGGALFAINCSITNNLGQGTFVDNLA
                     LNKGGALYTETNLSIKDNKGPIIIKQNRALNSDSLGGGIYSGNSLNIEGNSGAIQITS
                     NSSGSGGGIFSTQTLTISSNKKLIEISENSAFANNYGSNFNPGGGGLTTTFCTILNNR
                     EGVLFNNNQSQSNGGAIHAKSIIIKENGPVYFLNNTATRGGALLNLSAGSGNGSFILS
                     ADNGDIIFNNNTASKHALNPPYRNAIHSTPNMNLQIGARPGYRVLFYDPIEHELPSSF
                     PILFNFETGHTGTVLFSGEHVHQNFTDEMNFFSYLRNTSELRQGVLAVEDGAGLACYK
                     FFQRGGTLLLGQGAVITTAGTIPTPSSTPTTVGSTITLNHIAIDLPSILSFQAQAPKI
                     WIYPTKTGSTYTEDSNPTITISGTLTLRNSNNEDPYDSLDLSHSLEKVPLLYIVDVAA
                     QKINSSQLDLSTLNSGEHYGYQGIWSTYWVETTTITNPTSLLGANTKHKLLYANWSPL
                     GYRPHPERRGEFITNALWQSAYTALAGLHSLSSWDEEKGHAASLQGIGLLVHQKDKNG
                     FKGFRSHMTGYSATTEATSSQSPNFSLGFAQFFSKAKEHESQNSTSSHHYFSGMCIEN
                     TLFKEWIRLSVSLAYMFTSEHTHTMYQGLLEGNSQGSFHNHTLAGALSCVFLPQPHGE
                     SLQIYPFITALAIRGNLAAFQESGDHAREFSLHRPLTDVSLPVGIRASWKNHHRVPLV
                     WLTEISYRSTLYRQDPELHSKLLISQGTWTTQATPVTYNALGIKVKNTMQVFPKVTLS
                     LDYSADISSSTLSHYLNVASRMRF"
     gene            3196. .6000
                     /gene="pmp_16"
                     /note="CPn0467"
     CDS             3196. .6000
                     /gene="pmp_16"
                     /codon_start=1
                     /transl_table=11
                     /product="Polymorphic Outer Membrane Protein E Family"
                     /protein_id="AAD18609.1"
                     /db_xref="GI:4376752"
                     /translation="MFGMTPAVYSLQTDSLEKFALERDEEFRTSFPLLDSLSTLTGFS
                     PITTFVGNRHNSSQDIVLSNYKSIDNILLLWTSAGGAVSCNNFLLSNVEDHAFFSKNL
                     AIGTGGAIACQGACTITKNRGPLIFFSNRGLNNASTGGETRGGAIACNGDFTISQNQG
                     TFYFVNNSVNNWGGALSTNGHCRIQSNRAPLLFFNNTAPSGGGALRSENTTISDNTRP
                     IYFKNNCGNNGGAIQTSVTVAIKNNSGSVIFNNNTALSGSINSGNGSGGAIYTTNLSI
                     DDNPGTILFNNNYCIRDGGAICTQFLTIKNSGHVYFTNNQGNWGGALMLLQDSTCLLF
                     AEQGNIAFQNNEVFLTTFGRYNAIHCTPNSNLQLGANKGYTTAFFDPIEHQHPTTNPL
                     IFNPNANHQGTILFSSAYIPEASDYENNFISSSKNTSELRNGVLSIEDRAGWQFYKFT
                     QKGGILKLGHAASIATTANSETPSTSVGSQVIINNLAINLPSILAKGKAPTLWIRPLQ
                     SSAPFTEDNNPTITLSGPLTLLNEENRDPYDSIDLSEPLQNIHLLLSLSDVTARHINTD
                     NFHPESLNATEHYGYQGIWSPYWVETITTTNNASIETANTLYRALYANWTPLGYKVNP
                     EYQGDLATTPLWQSFHTMFSLLRSYNRTGDSDIERPFLEIQGIADGLFVHQNSIPGAP
                     GFRIQSTGYSLQASSETSLHQKISLGFAQFFTRTKEIGSSNNVSAHNTVSSLYVELPW
                     FQEAFATSTVLAYGYGDHHLHSLHPSHQEQAEGTCYSHTLAAAIGCSFPWQQKSYLHL
                     SPFVQAIAIRSHQTAFEEIGDNPRKFVSQKPFYNLTLPLGIQGKWQSKFHVPTEWTLE
                     LSYQPVLYQQNPQIGVTLLASGGSWDILGHNYVRNALGYKVHNQTALFRSLDLFLDYQ
                     GSVSSSTSTHHLQAGSTLKF"
     misc_feature    6195. .8026
                     /note="Polymorphic Outer Membrane Protein E Family;
                     contains frame shift"
     gene            6195. .6998
                     /gene="pmp_17.1"
                     /note="CPn0468"
     gene            6962. .8026
                     /gene="pmp_17.2"
                     /note="CPn0469"
     gene            7920. .9098
                     /gene="pmp_17.3"
                     /note="CPn0470"
     CDS             7920. .9098
                     /gene="pmp_17.3"
                     /codon_start=1
                     /transl_table=11
                     /product="Polymorphic Outer Membrane Protein (Frame-shift
                     with CPn0469)"
                     /protein_id="AAD18611.1"
                     /db_xref="GI:4376754"
                     /translation="MSLNLERISPLLYLLDVTAKKIDTSNLIVEAMNLDEHYGYQGIW
                     SPYWMETTTTSSTVPEQTNTNHRQLYVDWTPVGYRPNPERHGEFIANTLWQSAYNAL
                     LGIRILPPQNLKEHDLEASLQGLGLLINQHNREGRKGFRNHTTGYAATTSAKTAARHS
                     FSLGFAQMFSKTRERQSPSTTSSHNYFAGLRFDSLLFRDFISTGLSLGYSYGDHHMLC
                     HYTEILKGSSKAFFNNHTLVASLDCTFLPARITRTLELQPFISAIALRCSQASFQETG
                     DHIRKFHPKHPLTDLSSPIGFRSEWKTSHHIPMLWTTEISYVPTLYRKNPEMFTTLLI
                     SNGTWTTQATPVSYNSVAAKIKNTSQLFSRVTLSLDYSAQVSSSTVGQYLKAESHCTF
                     "
     gene            9127. .11967
                     /gene="pmp_18"
                     /note="CPn0471"
     CDS             9127. .11967
                     /gene="pmp_18"
                     /codon_start=1
                     /transl_table=11
                     /product="Polymorphic Outer Membrane Protein E/F Family"
                     /protein_id="AAD18610.1"
                     /db_xref="GI:4376753"
                     /translation="MQNNRSLSKSSFFVGALILGKTTILLNATPLSDYFDNQANQLTT
                     LFPLIDTLTNMTPYSHRATLFGVRDDTNQDIVLDHQNSIESWFENFSQDGGALSCKSL
                     AITNTKNQILFLNSFAIKRAGAMYVNGNFDLSENHGSIIFSGNLSFPNASNFADTCTG
                     GAVLCSKNVTISKNQGTAYFINNKAKSSGGAIQAAIINIKDNTGPCLFFNNAAGGTAG
                     GALFANACRIENNSQPIYFLNNQSGLGGAIRVHQECILTKNTGSVIFNNNFAMEADIS
                     ANHSSGGAIYCISCSIKDNPGIAAFDNNTAARDGGAICTQSLTIQDSGPVYFTNNQGT
                     WGGAIMLRQDGACTLFADQGDIIFYNNRHFKDTFSNHVSVNCTRNVSLTVGASQGHSA
                     TFYDPILQRYTIQNSIQKFNPNPEHLGTILFSSTYIPDTSTSRDDFISHFRNHIGLYN
                     GTLALEDRAEWKVYKFDQFGGTLRLGSRAVFSTTDEEQSSSSVGSVININNLAINLPS
                     ILGNRVAPKLWIRPTGSSAPYSEDNNPIINLSGPLSLLDDENLDPYDTADLAQPIAEV
                     PLLYLLDVTAKHINTDNFYPEGLNTTQHYGYQGVWSPYWIETITTSDTSSEDTVNTLH
                     RQLYGDWTPTGYKVNPENKGDIALSAFWQSFHNLFATLRYQTQQGQIAPTASGEATRL
                     FVHQNSNNDAKGFHMEATGYSLGTTSNTASNHSFGVNFSQLFSNLYESHSDNSVASHT
                     TTVALQINNPWLQERFSTSASLAYSYSNHHIKASGYSGKIQTEGKCYSTTLGAALSCS
                     LSLQWRSRPLHFTPFIQAIAVRSNQTAFQESGDKARKFSVHKPLYNLTVPLGIQSAWE
                     SKFRLPTYWNIELAYQPVLYQQNPEINVSLESSGSSWLLSGTTLARNAIAFKGRNQIF
                     IFPKLSVFLDYQGSVSSSTTTHYLHAGTTFKF"
BASE COUNT     3630 a    2929 c    2159 g    3365 t
ORIGIN

  Query Match             4.7%;  Score 136.4;  DB 1;  Length 12083;
  Best Local Similarity   45.9%;  Pred. No. 6.8e-21;
  Matches  762;  Conservative    0;  Mismatches  856;  Indels   42;  Gaps    7;

Qy    1253 CTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTG 1312
           || |   ||| |||    | ||  | ||    ||  |  |  |||||||   |||   |
Db    4367 CTCTAATCTTTAATCCCAATGCGAACCATCAGGGAACGATCTTATTTTCTTCAGCCTATA 4426

Qy    1313 TTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTC 1372
           |    ||   |   |  ||  |||||   |  |   |      ||  |  |
Db    4427 TCCCAGAAGCTTCTGACTACGAAAATAATTTCATTAGCAGCTCGAAAAATACCTCTGAAC 4486

Qy    1373 TCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACAC 1432
           |  | ||||||  | ||  | |||||| |||| |||          | |  ||   ||||| |
Db    4487 TTCGCAATGGTGTCCTCTCTATCGAGGATCGTGCGGGATGGCAATTCTATAAGTTCACTC 4546

Qy    1433 AAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTG 1492
           |||  || ||| |  ||    | ||| |||  ||   | |          |   ||   ||
Db    4547 AAAAAGGAGGTATCCTTAAATTAGGGCATGCGGCGAGTATTGCAACAACTGCCAACTCTG 4606

Qy    1493 CAGGAAATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTT 1552
             ||  ||||    | |||       | |   |||       || |
Db    4607 AGACTCCATCAACTAGTGTAGGCTCCCAGGTCATCATTAATAACCTTGCGATTAACCTCC 4666

Qy    1553 CCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCA 1612
           ||    |   |      |   |  ||||  ||||||  |    |||  | |    ||
Db    4667 CCTCGATCTTAGCAAAAGGAAAAGCTCCTACCTTGTGGATCCGTCCTCTACAATCTAGTG 4726

Qy    1613 ATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCA 1672
           |   |  ||||  || | |    ||||  || | |||   | |      ||  |||||
Db    4727 CTCCTTTCACAGAGGACAATAACCCTACAATTACTTTATCAGGTCCT---CTGACACTCT 4783
```

```
Qy   1673 TTGATGACTATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTGTCATCAC 1732
Db   4784 TAAATGAGGAAAACCGCGATCCCTACGACAGTATAGATCTCTCTGAGCCTTTACAAAACA 4843

Qy   1733 AGCCTATGCTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGATATGGATT 1792
Db   4844 TTCATCTTCTTTCTTTATCGGATGTAACAGCACGTCATATCAATACCGATAACTTTCATC 4903

Qy   1793 TTTCGGGACTAAATGTCCCT---CATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGG 1849
Db   4904 CTGAAAGCTTAAATGCGACTGAGCATTACGGTTATCAAGGCATCTGGTCTCCTTATTGGG 4963

Qy   1850 CAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAAGCCA 1909
Db   4964 ---------------TAGAGACGATAACAACAACAAATAACGCTTCTATAGAGACGGCAA 5008

Qy   1910 ATAGATTCCATAGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGCCCGA 1969
Db   5009 ACACCCTCTACAGAGCTCTGTATGCCAATTGGACTCCCTTAGGATATAAGGTCAATCCTG 5068

Qy   1970 AACACAGAAGTCCCCTCATAGCGAATACCTTATGG------GGGAATATGCTGCTTGCAA 2023
Db   5069 AATACCAAGGAGATCTTGCTACGACTCCCCTATGGCAATCCTTTCATACTATGTTCTCTC 5128

Qy   2024 CAGAAAGCTTAAAAAATAGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTA 2083
Db   5129 TATTAAGAAGTTATAATCGAACTGGTGATTCTGATATCGAGAGGCCTTTCTTAGAAATTC 5188

Qy   2084 CAGGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCC 2143
Db   5189 AAGGGATTGCCGACGGCCTCTTTGTTCATCAAAATAGCATCCCCGGGGCTCCAGGATTCC 5248

Qy   2144 ATATGCGCTCTTCCGGATACTC------TGCGGGGATGATAGCAGGGCAGACACACACCT 2197
Db   5249 GTATCCAATCTACAGGGTATTCCTTACAAGCATCCTCCGAAACTTCTTTACATCAGAAAA 5308

Qy   2198 TCTCATTGAAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTACGCAAAAAACAACG 2257
Db   5309 TCTCCTTAGGTTTTGCACAGTTCTTCACCCGCACTAAAGAAATCGGATCAAGCAACAACG 5368

Qy   2258 TATCTTCTAAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCT 2317
Db   5369 TCTCGGCTCACAATACAGTCTCTTCACTTTATGTTGAGCTTCCGTGGTTCCAAGAGGCCT 5428

Qy   2318 TGCTGACTAAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCACCATTTCTATA 2377
Db   5429 TTGCAACATCCACAGTGTTAGCGTATGGCTATGGGGACCATCACCTCCACAGCCTACATC 5488

Qy   2378 CCCAAGGAGAAAATCTAACATCTCAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTG 2437
Db   5489 CCTCACATCAAGAACAGGCA---GAAGGGACGTGTTATAGCCATACATTAGCAGCAGCTA 5545

Qy   2438 TTTTTTTTGATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCTCCCTTTT 2497
Db   5546 TCGGCTGTTCTTTCCCTTGGCAAC------AGAAATCCTATCTTCACCTCAGCCCGTTCG 5599

Qy   2498 TAGGTGCTCTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGC 2557
Db   5600 TTCAGGCAATTGCAATACGTTCTCACCAAACAGCGTTCGAAGAGATTGGTGACAATCCCC 5659

Qy   2558 GAAGCTTTTCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTA 2617
Db   5660 GAAAGTTTGTCTCTCAAAAGCCTTTCTATAATCTGACCTTACCTCTAGGAATCCAAGGAA 5719

Qy   2618 GCTTTATGAATGCTACCCAAAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCG 2677
Db   5720 AATGGCAGTCAAAATTCCACGTACCTACAGAATGGACTCTAGAACTTTCTTACCAACCGG 5779

Qy   2678 TTCTGTATAGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGT 2737
Db   5780 TACTCTATCAACAAAATCCCCAAATCGGTGTCACGCTACTTGCGAGCGGAGGTTCCTGGG 5839

Qy   2738 TTGGTAGTGGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAGCAAACAC 2797
Db   5840 ATATCCTAGGCCATAACTATGTTCGCAATGCTTTAGGGTACAAAGTCCACAATCAAACTG 5899

Qy   2798 AACCTTTGAGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCTTCAACCT 2857
Db   5900 CGCTCTTCCGTTCTCTCGATCTATTCTTGGATTACCAAGGATCGGTCTCCTCCTCGACAT 5959

Qy   2858 TCTGTAATTATCTCAATGGGGAAATTGCTCTGCGATTCTA 2897
Db   5960 CTACGCACCATCTCCAAGCAGGAAGTACCTTAAAATTCTA 5999
```

RESULT 11
AP002546

```
LOCUS       AP002546               300650 bp    DNA     linear   BCT 25-MAY-2002
DEFINITION  Chlamydophila pneumoniae J138 genomic DNA, complete sequence,
            section 2/4.
ACCESSION   AP002546 AB033782 AB033783 AB033784 AB033785 AB033800 AB033801
            AB033802 AB033803 AB033804 AB033805 AB033806 AB033807 AB033808
            AB033809 AB033810 AB033811 AB033812 AB033813 AB033814 AB033815
            AB036079 AB036080 AB036081 AB036082 AB038348 AB038349 BA000008
VERSION     AP002546.2  GI:10176692
KEYWORDS    .
SOURCE      Chlamydophila pneumoniae J138 (strain:J138) DNA.
  ORGANISM  Chlamydophila pneumoniae J138
            Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
REFERENCE   1
  AUTHORS   Shirai,M., Hirakawa,H., Ouchi,K., Tabuchi,M., Kishi,F., Kimoto,M.,
            Takeuchi,A., Nishida,J., Shibata,K., Fujinaga,R., Yoneda,H.,
            Matsushima,H., Tanaka,C., Furukawa,S., Miura,K., Nakazawa,A.,
            Ishii,K., Shiba,T., Hattori,M., Kuhara,S. and Nakazawa,T.
  TITLE     Comparison of outer membrane protein genes omp and pmp in the whole
            genome sequences of Chlamydia pneumoniae isolates from Japan and
            the United States
  JOURNAL   J. Infect. Dis. 181 Suppl 3, S524-S527 (2000)
  MEDLINE   20298986
REFERENCE   2
  AUTHORS   Shirai,M., Hirakawa,H., Kimoto,M., Tabuchi,M., Kishi,F., Ouchi,K.,
            Shiba,T., Ishii,K., Hattori,M., Kuhara,S. and Nakazawa,T.
  TITLE     Comparison of whole genome sequences of Chlamydia pneumoniae J138
            from Japan and CWL029 from USA
  JOURNAL   Nucleic Acids Res. 28 (12), 2311-2314 (2000)
  MEDLINE   20330349
REFERENCE   3  (bases 1 to 300650)
  AUTHORS   Shirai,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2000) Mutsunori Shirai, Yamaguchi University
            School of Medicine, Department of Microbiology; 1-1-1
            Minamikogushi, Ube, Yamaguchi 755-8505, Japan
            (E-mail:mshirai@po.cc.yamaguchi-u.ac.jp, Tel:81-836-22-2227,
            Fax:81-836-22-2415)
COMMENT     On or before Sep 15, 2000 this sequence version replaced
            gi:6172290, gi:6172292, gi:6172294, gi:6172296, gi:6172326,
            gi:6172328, gi:6172330, gi:6172332, gi:6172334, gi:6172336,
            gi:6172338, gi:6172382, gi:6172384, gi:6172386, gi:6172388,
            gi:6172390, gi:6172392, gi:6172394, gi:6174666, gi:6174668,
            gi:6635174, gi:6635176, gi:6635178, gi:6635180, gi:8547433,
            gi:8547438, gi:8978640.
FEATURES             Location/Qualifiers
     source          1. .300650
                     /organism="Chlamydophila pneumoniae J138"
                     /strain="J138"
                     /db_xref="taxon:138677"
                     /note="synonym:Chlamydia pneumoniae (strain J138)"
     gene            144. .935
                     /gene="CPj0267"
     CDS             144. .935
                     /gene="CPj0267"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAA98477.1"
```

/db_xref="GI:8978641"
/translation="MSIMSLNKTNALLNQPEPAVCLNAWDPKYINQDRKTFACTVTLL
VIATLMILTTGVIVLLAMGSPGLSVLVSTIIGTSVTTLGTALFIIGLVKLIKKSLAWI
QYQKYFQEVVKQKYEPFSIPKNDNVHKLTSCLPSPLDIESPSPEASTPVSKLRIACSG
VAIVLGVTLLIGAVVSVFFCTGYLQLALCVGFACLGTALFVGGLAGLRTHSLIAQGIM
YLYLTYYLSSALEERNETVKDQRNEINTYLTEECRQQKREKALLE"

gene 957. .1343
/gene="CPj0268"

CDS 957. .1343
/gene="CPj0268"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAA98478.1"
/db_xref="GI:8978642"
/translation="MSQCQSSSTSTWEWMKSFVPNWKNPTPPLSPIPSEDEFILAYEP
FVLPKTDPENAQANPPGTSTPNVENGIDDLNPLLGQPNEQNNANNPGTSGSNPTSLPA
PERLPETEENSQEEEQGSQNNEDLIG"

gene complement(1495. .2472)
/gene="CPj0269"

CDS complement(1495. .2472)
/gene="CPj0269"
/codon_start=1
/transl_table=11
/product="dipeptidase"
/protein_id="BAA98479.1"
/db_xref="GI:8978643"
/translation="MTIDMHCDLLSHPHFCRKDPAVRCSPEQLLSGGVRQQVCAIFVP
HSRGEPNCDKQNSLFFSLPNQYPDIGLLSYEEEENGSSSQKKSLSLIRSIENASALGD
DTAPLGTLLAKLIHLTKQGPLAYLGIVWKGDNRFGGGTEAPKRLSNDGKVLLDIMYEL
GVPIDLSHCSDKLAEDILDYTADKLPNLAVIASHSNFRSVLDHRRNLVDAHAKEIVRR
KGVIGLNLVRSYVGDSLGDLEKHVLHAENLGILSSIVLGSDFFYANEDENFFFNECSS
AEAHPVLNQLIHRIFSKGKAESILSSRAEKFLKQVIVEQVNPKITDVKL"

gene complement(2487. .3347)
/gene="ywlC"

CDS complement(2487. .3347)
/gene="ywlC"
/codon_start=1
/transl_table=11
/product="SuA5 superfamily-related protein"
/protein_id="BAA98480.1"
/db_xref="GI:8978644"
/translation="MPDKKAQITFSLPEVMSAIHQGKIVALPTDTVYGFVLSLYASEA
EERLYALKDREPSKAFALYVNSIEDIENISGYPLSPTAKKLAQLFPGAITLVVKHRNP
RFPKETLAFRIVDHSVVREIVDHCGTLIGTSANLSEFPSALTAQEIFADFADHDLCIF
DGPCSHGLESTVVASDPLYIYREGLISRSVIENIAGTEAKIFHRTSHAFSKHIKIYTV
KNQEQLVSFLSGSLDFKGVVCEHPKPKNFYTRLREALKKKTPSIVFIYDINTSDYPEL
FPFLSPYYIE"

gene 3656. .4387
/gene="CPj0271"

CDS 3656. .4387
/gene="CPj0271"
/codon_start=1
/transl_table=11
/product="lysophospholipase esterase"
/protein_id="BAA98481.1"
/db_xref="GI:8978645"
/translation="MTDYSFFRRKIGNIEAIECPGNPQDPIIILCHGYGSLADNLTFF
PSICSFSKLRPTWIFPNGILPLENDFRGSRACFPLNVLLLQELSRLYANGVGNLQEKY
DELFDVDLETPKEALEELILNLNRPYNEIIIGGFSQGAILATHLVLTSQNPYAGALIF
AGARLFNQGWEEGLKQCAQVPFLQSHGYEDEILPYHLGAHLNDLLLTKLNGQFVSFHG
GHEIPSVVFQKMQVTVPNWIDPARG"

gene complement(4359. .5255)
/gene="dnaX_2"

CDS complement(4359. .5255)
/gene="dnaX_2"
/codon_start=1
/transl_table=11
/product="DNA polymerase III gamma and tau"
/protein_id="BAA98482.1"
/db_xref="GI:8978646"
/translation="MHLEEENQGWEALLRKVYHQEVPPAILLHGFTLPVLQDKAEQLA
SEILLSSSPGSEHKVSQKIHPDIYQFFPEGKGRLHSIDLPRGIKKQIYISPFEANYKI
YIIHEADRMTLAAISAFLKVFEEPPKHAVIILTTAKVQRLPKTIISRSLSIFIERGEK
ILCSKETFSYLFRYAQCEIPVTEVSQIIKESSETDKQVLRDKVQRFMEVLLELYRDRY
TLNLGLKASALNYPEHVKEILQLPLLPLDKVLLIVESACRSLNNSSSAASVLEWVAIQ
LVSLQYKEKELVSVSPGQDLSN"

gene complement(5246. .5866)
/gene="tdk"

CDS complement(5246. .5866)
/gene="tdk"
/codon_start=1
/transl_table=11
/product="thymidylate kinase"
/protein_id="BAA98483.1"
/db_xref="GI:8978647"
/translation="MFIVIEGGEGSGKSSLAKALGDQLVAQDRKVLLTREPGGCLIGE
RLRDLILEPPHLELSRCCELFLFLGSRAQHIQEVIIPALRDGYIVICERFHDSTIVYQ
GIAEGLGADFVADLCSKVVGPTPFLPNFVLLLDIPADIGLQRKHRQKVFDKFEKKPLS
YHNRIREGFLSLASADPSRYLVLDARESLASLIDKVMLHTQLGLCT"

gene complement(5871. .8375)
/gene="gyrA_1"

CDS complement(5871. .8375)
/gene="gyrA_1"
/codon_start=1
/transl_table=11
/product="DNA gyrase subunit A"
/protein_id="BAA98484.1"
/db_xref="GI:8978648"
/translation="MFNKDEIIVPKNLEEEMKESYLRYSMSVIISRALPDIRDGLKPS
QRRVLYAMKQLSLSPGAKHRKCAKICGDTSGDYHPHGESVIYPTLVRMAQNWAMRYPL
VDGQGNFGSIDGDPPAAMRYTEARLTHSAMYLMEDLDKDTVDIVPNYDETKHEPVVFP
SKFPNLLCNGSSGIAVGMATNIPPHNLGELIEATLLLLANPQASVDEILQVMPGPDFP
TGGIICGSEGIRSTYTTGRGKIKVRARLHVEENEDKHRESIIITEMPYNVNKSRLIEQ
IANLVNEKTLAGISDVRDESDKDGIRVVLEIKKGESSEIIINRLYKFTDVQVTFGANM
LALDKNLPRTMSIHRMISAWIRHRKEVIRRRTRYELNKAETRAHVLEGYLKALSCLDA
LVKTIRESGNKEHAKERIIESFGFTEPQALAILELRLYQLTGLEAEKIQKEYEELLNK
IAYYKQVLSDEGLVKDIIRNELQDLLKHHKVARRTTIEFDADDIRDIEDIITNESVII
TISGDDYVKRMPVKVFKEQRRGGHGVTGFDMKKGAGFLKAVYSAFTKDYLLIFTNFGQ
CYWLKVWQLPEGERRAKGKPIINFLEGIRPGEELAAILNIKNFDNAGFLFLATKRGVV
KKVSLDAFSNPRKKGIRALEIDEGDELIAACHIVSDEEKVMLFTHLGMAVRFPHEKVR
PMGRTARGVRGVSLKNEEDKVVSCQIVTENQSVLIVCDQGFGKRSLVEDFRETNRGGV
GVRSILINERNGNVLGAIPVTDHDSILLMSSQGQAIRINMQDVRVMGRSTQGVRLVHL
KEGDALVSMEKLSSNENDDEVLSGSEEECSDTVSLR"

gene complement(8391. .10808)
/gene="gyrB_1"

CDS complement(8391. .10808)
/gene="gyrB_1"
/codon_start=1
/transl_table=11
/product="DNA gyrase subunit B"
/protein_id="BAA98485.1"
/db_xref="GI:8978649"
/translation="MDPKEKNYDASAITVLEGLQAVRERPGMYIGDTGITGLHHLVYE
VVDNSIDEAMAGYCSRIDVRILEDGGIVIVDNGRGIPIEVHERESAKQGREVSALEVV
LTVLHAGGKFDKDSYKVSGGLHGVGVSCVNALSEKLVATVFKDKKCYQMEFSRGIPVT
PLQYVSVSDRQGTEIVFYPDPKIFSTCTFDRSILMKRLRELAFLNRGITIVFEDDRDV
SFDKVTFFYEGGIQSFVSYLNQNKESLFSEPIYICGTRVGDDGEIEFEAALQWNSGYS
ELVYSYANNIPTRQGGTHLTGFSTALTRVINTYIKAHNLAKNNKLALTGEDIREGLTA
VISVKVPNPQFEGQTKQKLGNSDVSSVAQQVVGEALTIFFEENPQIARMIVDKVFVAA
QAREAAKKARELTLRKSALDSARLPGKLIDCLEKDPEKCEMYIVEGDSAGGSAKQGRD
RRFQAILPIRGKILNVEKARLQKIFQNQEIGTIIAALGCGIGADNFNLSKLRYRRIII
MTDADVDGSHIRTLLLTFFYRHMTALIENECVYIAQPPLYKVSKKKDFRYILSEKEMD
SYLLMLGTNESSILFKSTERELRGEALESFINVILDVESFINTLEKKAIPFSEFLEMY
KEGIGYPLYYLAPATGMQGGRYLYSDEEKEEALAQEETHKFKIIELYKVAVFVDIQNQ
LKEYGLDISSYLIPQKNEIVIGNEDSPSCNYSCYTLEEVINYLKNLGRKGIEIQRYKG
LGEMNADQLWDTTMNPEQRTLIHVSLKDAVEADHIFTMLMGEEVPPRREFIESHALSI
RINNLDI"

gene complement(10812. .11159)
/gene="CPj0276"

CDS complement(10812. .11159)
/gene="CPj0276"

Query Match 4.7%; Score 136.4; DB 1; Length 300650;
Best Local Similarity 45.9%; Pred. No. 4.5e-21;
Matches 762; Conservative 0; Mismatches 856; Indels 42; Gaps 7;

```
Qy    1253 CTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTG 1312
Db 237523 CTCTAATCTTTAATCCCAATGCGAACCATCAGGGAACGATCTTATTTTCTTCAGCCTATA 237582

Qy    1313 TTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTC 1372
Db 237583 TCCCAGAAGCTTCTGACTACGAAAATAATTTCATTAGCAGCTCGAAAAATACCTCTGAAC 237642

Qy    1373 TCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACAC 1432
Db 237643 TTCGCAATGGTGTCCTCTCTATCGAGGATCGTGCGGGATGGCAATTCTATAAGTTCACTC 237702

Qy    1433 AAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTG 1492
Db 237703 AAAAAGGAGGTATCCTTAAATTAGGGCATGCGGCGAGTATTGCAACAACTGCCAACTCTG 237762

Qy    1493 CAGGAAATTCTGCTAGCAATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTT 1552
Db 237763 AGACTCCATCAACTAGTGTAGGCTCCCAGGTCATCATTAATAACCTTGCGATTAACCTCC 237822

Qy    1553 CCATTCTGAAAAGTGGTGCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCA 1612
Db 237823 CCTCGATCTTAGCAAAAGGAAAAGCTCCTACCTTGTGGATCCGTCCTCTACAATCTAGTG 237882

Qy    1613 ATAACTATACAGCAGATACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCA 1672
Db 237883 CTCCTTTCACAGAGGACAATAACCCTACAATTACTTTATCAGGTCCT---CTGACACTCT 237939

Qy    1673 TTGATGACTATGGGAATTCTCCTTATGAATCCACAGATCTAACCCATGCTCTGTCATCAC 1732
Db 237940 TAAATGAGGAAAACCGCGATCCCTACGACAGTATAGATCTCTCTGAGCCTTTACAAAACA 237999

Qy    1733 AGCCTATGCTATCTATTTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGATATGGATT 1792
Db 238000 TTCATCTTCTTTCTTTATCGGATGTAACAGCACGTCATATCAATACCGATAACTTTCATC 238059

Qy    1793 TTTCGGGACTAAATGTCCCT---CATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGG 1849
Db 238060 CTGAAAGCTTAAATGCGACTGAGCATTACGGTTATCAAGGCATCTGGTCTCCTTATTGGG 238119

Qy    1850 CAAAAACTCAAGATCCAGAACCAGCATCTTCAGCAACAATCACAGATCCACAAAAAGCCA 1909
Db 238120 --------------TAGAGACGATAACAACAACAAATAACGCTTCTATAGAGACGGCAA 238164

Qy    1910 ATAGATTCCATAGAACCTTATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGCCCGA 1969
Db 238165 ACACCCTCTACAGAGCTCTGTATGCCAATTGGACTCCCTTAGGATATAAGGTCAATCCTG 238224

Qy    1970 AACACAGAAGTCCCCTCATAGCGAATACCTTATGG------GGGAATATGCTGCTTGCAA 2023
Db 238225 AATACCAAGGAGATCTTGCTACGACTCCCCTATGGCAATCCTTTCATACTATGTTCTCTC 238284

Qy    2024 CAGAAAGCTTAAAAAATAGTGCAGAACTGACACCTAGTGATCATCCTTTCTGGGGAATTA 2083
Db 238285 TATTAAGAAGTTATAATCGAACTGGTGATTCTGATATCGAGAGGCCTTTCTTAGAAATTC 238344

Qy    2084 CAGGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCGAGAAAATCATCCTGGATTCC 2143
Db 238345 AAGGGATTGCCGACGGCCTCTTTGTTCATCAAAATAGCATCCCCGGGGCTCCAGGATTCC 238404

Qy    2144 ATATGCGCTCTTCCGGATACTC------TGCGGGGATGATAGCAGGGCAGACACACACCT 2197
Db 238405 GTATCCAATCTACAGGGTATTCCTTACAAGCATCCTCCGAAACTTCTTTACATCAGAAAA 238464

Qy    2198 TCTCATTGAAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTACGCAAAAAACAACG 2257
Db 238465 TCTCCTTAGGTTTTGCACAGTTCTTCACCCGCACTAAAGAAATCGGATCAAGCAACAACG 238524

Qy    2258 TATCTTCTAAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCT 2317
Db 238525 TCTCGGCTCACAATACAGTCTCTTCACTTTATGTTGAGCTTCCGTGGTTCCAAGAGGCCT 238584

Qy    2318 TGCTGACTAAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTGTCACCATTTCTATA 2377
Db 238585 TTGCAACATCCACAGTGTTAGCGTATGGCTATGGGGACCATCACCTCCACAGCCTACATC 238644

Qy    2378 CCCAAGGAGAAAATCTAACATCTCAAGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTG 2437
Db 238645 CCTCACATCAAGAACAGGCA---GAAGGGACGTGTTATAGCCATACATTAGCAGCAGCTA 238701

Qy    2438 TTTTTTTTGATCTCCCTATGAAACCCTTTGGATCAACGCATATACTGACAGCTCCCTTTT 2497
Db 238702 TCGGCTGTTCTTTCCCTTGGCAAC------AGAAATCCTATCTTCACCTCAGCCCGTTCG 238755

Qy    2498 TAGGTGCTCTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGC 2557
Db 238756 TTCAGGCAATTGCAATACGTTCTCACCAAACAGCGTTCGAAGAGATTGGTGACAATCCCC 238815

Qy    2558 GAAGCTTTTCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTA 2617
Db 238816 GAAAGTTTGTCTCTCAAAAGCCTTTCTATAATCTGACCTTACCTCTAGGAATCCAAGGAA 238875

Qy    2618 GCTTTATGAATGCTACCCAAAGACCTCAAGCCTGGACTGTAGAATTGGCATACCAACCCG 2677
Db 238876 AATGGCAGTCAAAATTCCACGTACCTACAGAATGGACTCTAGAACTTTCTTACCAACCGG 238935

Qy    2678 TTCTGTATAGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGT 2737
Db 238936 TACTCTATCAACAAAATCCCCAAATCGGTGTCACGCTACTTGCGAGCGGAGGTTCCTGGG 238995

Qy    2738 TTGGTAGTGGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAATCTCACAGCAAACAC 2797
Db 238996 ATATCCTAGGCCATAACTATGTTCGCAATGCTTTAGGGTACAAAGTCCACAATCAAACTG 239055

Qy    2798 AACCTTTGAGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTACTCCTCTTCAACCT 2857
Db 239056 CGCTCTTCCGTTCTCTCGATCTATTCTTGGATTACCAAGGATCGGTCTCCTCCTCGACAT 239115

Qy    2858 TCTGTAATTATCTCAATGGGGAAATTGCTCTGCGATTCTA 2897
Db 239116 CTACGCACCATCTCCAAGCAGGAAGTACCTTAAAATTCTA 239155
```

```
RESULT 12
AE002189/c
LOCUS       AE002189               10236 bp    DNA     linear   BCT 30-MAY-2000
DEFINITION  Chlamydophila pneumoniae AR39, section 22 of 94 of the complete
            genome.
ACCESSION   AE002189 AE002161
VERSION     AE002189.2  GI:8163402
KEYWORDS    .
SOURCE      Chlamydophila pneumoniae AR39.
  ORGANISM  Chlamydophila pneumoniae AR39
            Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
REFERENCE   1  (bases 1 to 10236)
  AUTHORS   Read,T.D., Brunham,R.C., Shen,C., Gill,S.R., Heidelberg,J.F.,
            White,O., Hickey,E.K., Peterson,J., Umayam,L.A., Utterback,T.,
            Berry,K., Bass,S., Linher,K., Weidman,J., Khouri,H., Craven,B.,
            Bowman,C., Dodson,R., Gwinn,M., Nelson,W., DeBoy,R., Kolonay,J.,
            McClarty,G., Salzberg,S.L., Eisen,J. and Fraser,C.M.
  TITLE     Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
            pneumoniae AR39
  JOURNAL   Nucleic Acids Res. 28 (6), 1397-1406 (2000)
  MEDLINE   20150255
   PUBMED   10684935
REFERENCE   2  (bases 1 to 10236)
  AUTHORS   Read,T.D., Brunham,R.C., Shen,C., Gill,S.R., Heidelberg,J.F.,
            White,O., Hickey,E.K., Peterson,J., Umayam,L.A., Utterback,T.,
            Berry,K., Bass,S., Linher,K., Weidman,J., Khouri,H., Craven,B.,
            Bowman,C., Dodson,R., Gwinn,M., Nelson,W., DeBoy,R., Kolonay,J.,
            McClarty,G., Salzberg,S.L., Eisen,J. and Fraser,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAR-2000) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
COMMENT     On Jun 1, 2000 this sequence version replaced gi:7189205.
```

```
FEATURES             Location/Qualifiers
     source          1. .10236
                     /organism="Chlamydophila pneumoniae AR39"
                     /strain="AR39"
                     /db_xref="taxon:115711"
                     /note="synonym: Chlamydia pneumoniae AR39"
     gene            135. .1661
                     /gene="CP0281"
     CDS             135. .1661
                     /gene="CP0281"
                     /note="hypothetical protein; identified by Glimmer2;
                     putative"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAF38139.1"
                     /db_xref="GI:7189206"
                     /translation="MAVGGVGGSRSPSPIPPNRRNSEDGKVSPKDNLGEHTVSSSDSS
                     LASQGPTIEERKAQLGGTDKIPLPSVKEPGDSPTSGRSGVLQRIWKGVKGVFKKTPQA
                     RPEVSSPRLPSHVQHGQRLPGLEGFRDRIQKRSENPEADLGKMKRSYSDGDLDRVGHD
                     SNEDSTEDSRSEGGEPSSKSSSFLSGVRGAVSKVHGALGDIKGKFQRSASEDDLTTQG
                     EDSAGDTVKERRSEEAEASSKSSSFLSGVRGATSTVQGALGDAKEKVSAFGEQAAGAI
                     RSAPGNIRTRFQRSSSEGDLSNVNKAAKHLRKALENLEKVAPEQVSPEVASRVQSLLA
                     RMEQLTHQEPPTVEDLITFVESNVGSDSVEYASIVPQDGSQAPAETAEAPETGGVEGS
                     AAQGAWKALRDFVVSIFQAVASFFRAIASRLSSARRESAVDDLASESNTQWFVEQEGV
                     SNPSAAPSLSFAEEIARRAAEMSNRNAQSLEKLESGNVTDPVIQQGLGLARSFAPEGQ
                     "
     gene            1823. .4150
                     /gene="CP0282"
     CDS             1823. .4150
                     /gene="CP0282"
                     /note="hypothetical protein; identified by Glimmer2;
                     putative"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAF38140.1"
                     /db_xref="GI:7189207"
                     /translation="MASGIGGSSGLGKIPPKDNGDRSRSPSPKGELGSHEISLPPQEH
                     GEEGASGSSHIHSSSSFLPEDQESQSSSSAASSPGFFSRVRSGVDRALKSFGNFFSAE
                     STSQARETRQAFVRLSKTITADERRDVDSSSAAATEARVAEDASVSGENPSQGVPETS
                     SGPEPQRLFSLPSVKKQSGLGRLVQTVRDRIVLPSGAPPTDSEPLSLYELNLRLSSLR
                     QELSDIQSNDQLTPEEKAEATVTIQQLIQITEFQCGYMEATQSSVSLAEARFKGVETS
                     DEINSLCSELTDPELQELMSDGDSLQNLLDETADDLEAALSHARLSFSLDDNPTPIDN
                     NPTLISQEEPIYEEIGGAADPQRTRENWSTRLWNQIREALVSLLGMILSILGSILHRL
                     RIARHAAAEAVGRCCTCRGEECTSSEEDSMSVGSPSEIDETERTGSPHDVPRRNGSPR
                     EDSPLMNALVGWAHKHGAKTKESSESSTPEISISAPIVRGWSQDSSVSFIVMEDDHIF
                     YDVPRRKDGIYDVPSSPRWSPARELEEDVFGDYEVPITSAEPSKDKNIYMTPRLATPA
                     IYDLPSRPGSSGSSRSPSSDRVRSSSPNRRGVPLPPVPSPAMSEEGSIYEDMSGASGA
                     GESDYEDMSRSPSPRGDLDEPIYANTPEDNPFTQRNIDRILQERSGGASASPVEPIYD
                     EIPWIHGRPPATLPRPENTLTNVSLRVSPGFGPEVRAALLSESVSAVMVEAESIVPPT
                     EPGDGESEYLEPLGGLVATTKILLQKGWPRGESNA"
     gene            complement(4324. .7164)
                     /gene="CP0283"
     CDS             complement(4324. .7164)
                     /gene="CP0283"
                     /note="similar to GP:4376753; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="polymorphic membrane protein E/F family"
                     /protein_id="AAF38141.1"
                     /db_xref="GI:7189208"
                     /translation="MQNNRSLSKSSFFVGALILGKTTILLNATPLSDYFDNQANQLTT
                     LFPLIDTLTNMTPYSHRATLFGVRDDTNQDIVLDHQNSIESWFENFSQDGGALSCKSL
                     AITNTKNQILFLNSFAIKRAGAMYVNGNFDLSENHGSIIFSGNLSFPNASNFADTCTG
                     GAVLCSKNVTISKNQGTAYFINNKAKSSGGAIQAAIINIKDNTGPCLFFNNAAGGTAG
                     GALFANACRIENNSQPIYFLNNQSGLGGAIRVHQECILTKNTGSVIFNNNFAMEADIS
                     ANHSSGGAIYCISCSIKDNPGIAAFDNNTAARDGGAICTQSLTIQDSGPVYFTNNQGT
                     WGGAIMLRQDGACTLFADQGDIIFYNNRHFKDTFSNHVSVNCTRNVSLTVGASQGHSA
                     TFYDPILQRYTIQNSIQKFNPNPEHLGTILFSSAYIPDTSTSRDDFISHFRNHIGLYN
                     GTLALEDRAEWKVYKFDQFGGTLRLGSRAVFSTTDEEQSSSSVGSVININNLAINLPS
                     ILGNRVAPKLWIRPTGSSAPYSEDNNPIINLSGPLSLLLDDENLDPYDTADLAQPIAEV
                     PLLYLLDVTAKHINTDNFYPEGLNTTQHYGYQGVWSPYWIETITTSDTSSEDTVNTLH
                     RQLYGDWTPTGYKVNPENKGDIALSAFWQSFHNLFATLRYQTQQGQIAPTASGEATRL
                     FVHQNSNNDAKGFHMEATGYSLGTTSNTASNHSFGVNFSQLFSNLYESHSDNSVASHT
                     TTVALQINNPWLQERFSTSASLAYSYSNHHIKASGYSGKIQTEGKCYSTTLGAALSCS
                     LSLQWRSRPLHFTPFIQAIAVRSNQTAFQESGDKARKFSVHKPLYNLTVPLGIQSAWE
                     SKFRLPTYWNIELAYQPVLYQQNPEVNVSLESSGSSWLLSGTTLARNAIAFKGRNQIF
                     IFPKLSVFLDYQGSVSSSTTTHYLHAGTTFKF"
     gene            complement(7193. .10096)
                     /gene="CP0284"
                     /note="This region contains a gene with one or more
                     premature stops or frameshifts, and is not the result of a
                     sequencing artifact; similar to GP:4376751; identified by
                     sequence similarity; putative;polymorphic membrane protein
                     E/F family, degenerate"
BASE COUNT     2874 a   1885 c   2466 g   3011 t
ORIGIN

  Query Match             4.3%;  Score 124.2;  DB 1;  Length 10236;
  Best Local Similarity   44.3%;  Pred. No. 4.5e-18;
  Matches  936;  Conservative    0;  Mismatches 1118;  Indels   61;  Gaps    8;

Qy     790 AACAATCGCTGCTTTAAAAATGTAGAAACAGCTTCTTCAGAAGCTTCTGATGGAGGAGCA 849
           | ||||| || |          | | ||||| ||   | ||   | |   || || |
Db    9254 AGCAATCACTCCGCATCCTCAATTAACACAGCATCAGGAAAACTATATCCCGGTGGTGGC 9195

Qy     850 ATTAAAGTAACTACTCGCCTAGATGTTACAGGCAATCGTGGTAGGATCTTTTTTAGTGAC 909
             ||   |  |||   | |||   ||     |||||      |  |  | ||||   |
Db    9194 GGTATCATGTGTACCTCCTTAGTCATTGAGAACAATCCCAAAGGTCTTATCTTTAACAAT 9135

Qy     910 AATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTACCCTAGTGGATAAT 969
           || |   ||  |  |  ||||||||||||  || |       | | | |     ||||
Db    9134 AAAACGGCAGCACTTAGCGGCGGAGCTATACACACGAGATCTTTCATCTTCCAAAATAAC 9075

Qy     970 GGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCGCTATCTATATAGAC 1029
           || || ||    ||||| || ||    || | |      ||||     ||| || |
Db    9074 GGTCCGACAGCATTTATTAATAACTCTGCGACTTCAGGAGGGGCTCTCATCAATCTTTCT 9015

Qy    1030 GGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTTTTAATGAAAATATT 1089
           || |   | |   |          |  |  || |  |||             ||  |
Db    9014 GGTATAGGAAGTACTCCTCAAAATTTCTTCCTCTCTGCA-----------GACTACGGC 8967

Qy    1090 GTGACTAATGTAACTAATGCAAATGGTACCAGTACGTCAGCTAATCCTCCTAGAAGAAAT 1149
           |  | |    | |  |||   |     ||   | | ||  || | ||       ||||||
Db    8966 GATATTCTATTTAACAATAATACAATCACATCTTCTTCTCCTCAACCCGGATATAGAAAT 8907

Qy    1150 GCAATAACAGTAGCAAGCTCCTCTGGTGAAATTCTATTAGGAGCAGGGAGTAGCCAAAAT 1209
           |||  |    |  ||          |     ||  |||  || |  |  ||    ||| |
Db    8906 GCACTCTATGCTGCTCCGGGGATTAACTTAAAACTAGGAGCAAGACAGGGTTATAAAATT 8847

Qy    1210 TTAATTTTTTATGATCCTATTGAAGTTAGCAATGCAGGGGTCTCTGTGTCCTTCAATAAG 1269
           |  ||| | ||  |     | | | |   |  ||   |   || |    || ||| |
Db    8846 CTCTTTTATGATCCTATAGATCACGATCAGACGACAACAGATCCTATAGTATTTAATTAT 8787

Qy    1270 GAAGCTGATCAAACAGGCTCTGTAGTATTTTCAGGAGCTACTGTTAATTCTGCAGATTTT 1329
           ||| |  ||||    ||| | ||  | ||||| |||   | |||  |||||   |
Db    8786 GAACCCCATCACCTTGGCACCGTGTTGTTTTCCGGAATCAATGTAGATTCTAACGCAACA 8727

Qy    1330 CATCAACGCAATTTACAAACAAAAACACCTGCACCCCTTACTCTCAGTAATGGTTTTCTA 1389
             ||| |   || || | | ||||    ||    |       ||    |  ||| | ||
Db    8726 AATCCATTGAACTTCCTATCAAAATTTTCTAACTCTTCACGACTTGAAAGGGGTGTGCTC 8667

Qy    1390 TGTATCGAAGATCATGCTCAGCTTACAGTGAATCGATTCACACAAACTGGGGGTGTTGTT 1449
           |||  |||||||  |||    || |    ||      |  | ||||||||||  || |
Db    8666 GCTATTGAAGATCGGGCTGCTATTTCTTGCAAAACCCTATCGCAAACTGGGGGCATTCTA 8607

Qy    1450 TCTCTTGGGAATGGAGCAGTTCTGAGTTGCTATAAAAATGGTGCAGGAAATTCTGCTAGC 1509
             | | || || | ||||  |   | ||       |  || ||   | ||||
Db    8606 CGTTTAGGAAACGCAGCATTAATCAGG------ACGAAAGGCCCGGGAA---------- 8564

Qy    1510 AATGCCTCTATAACACTGAAGCATATTGGATTGAATCTTTCTTCCATTCTGAAAAGTGGT 1569
             ||| ||||   | ||    || |   | |||||| |||| ||| |  ||   |
Db    8563 ----GCTCCATAAATTTTAATGCAATCGCGATCAATCTTCCTTCTATTTTACAATCAGAA 8508
```

```
Qy   1570 GCTGAGATTCCTTTATTGTGGGTAGAGCCTACAAATAACAGCAATAACTATACAGCAGAT 1629
Db   8507 GCCTCAGCTCCAAAGTTCTGGATTTATCCTACATTAACAGGATCCACCTATTCTGAAGAC 8448

Qy   1630 ACTGCAGCTACCTTTTCATTAAGTGATGTAAAACTCTCACTCATTGATGACTATGGGAAT 1689
Db   8447 ACTTCTTCTACTATCACTCTCTCAGGACCCTTGACTTTTCTAAACGATGAAAATGAAAAC 8388

Qy   1690 TCTCCTTAT-GAATCCACAGATCTAACCCATGCTCTGTCATCACAGCCTATGCTATCTAT 1748
Db   8387 CCCTATGATAGCTTAGATCTCTCTGAACCTCGAAAGGATATCCCCCCTCCTCTACCTCCT 8328

Qy   1749 TTCTGAGGCTAGTGATAACCAGCTAAGATCTGATGATATGGATTTTTCGGGACTAAATGT 1808
Db   8327 CGATGTGACTGCAAAAAAAATCGATACTTCGAATCTCATTGTAGAAGCCATGAACTTAGA 8268

Qy   1809 CCCTCATTATGGATGGCAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAAGATCCAGA 1868
Db   8267 TGAGCACTATGGATATCAGGGAATCTGGTCTCCCTATTGGATGGAAACT----------- 8219

Qy   1869 ACCAGCATCTTCAGCAACAATCACAGATCCACAAAAAGCCAATAGATTCCATAGAACCTT 1928
Db   8218 ----ACGACTACAACAAGCTCTACAGTACCGGAACAGACCAATACAAACCACAGGCAGCT 8163

Qy   1929 ATTACTGACTTGGCTTCCTGCTGGGTATGTTCCTAGCCCGAAACACAGAAGTCCCCTCAT 1988
Db   8162 CTACGTAGACTGGACTCCTGTAGGATACCGCCCTAACCCGGAACGTCACGGAGAATTTAT 8103

Qy   1989 AGCGAATACCTTATGGGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAATAGTGCAGA 2048
Db   8102 TGCTAATACCTTATGGCAGTCTGCCTATAACGCTCTGTTAGGAATCCGCATCTTACCTCC 8043

Qy   2049 ACTGACACCTAGTGATCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGCATGATGGT 2108
Db   8042 ACAAAACCTCAAAGAGCA---TGACCTTGAAGCCTCTCTGCAAGGACTCGGGCTTCTAAT 7986

Qy   2109 TTACCAAGATCCTCGAGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGATACTCTGC 2168
Db   7985 TAACCAACATAATCGCGAGGGACGCAAAGGCTTCCGAAACCATACTACGGGCTATGCAGC 7926

Qy   2169 GGGGATGATAGCAGGGCAGAC------ACACACCTTCTCATTGAAATTCAGTCAGACCTA 2222
Db   7925 AACAACCTCAGCAAAAACTGCAGCACGACATAGTTTCTCTTTAGGATTCGCACAAATGTT 7866

Qy   2223 CACCAAACTCAATGAGCGTTACGCAAAAAACAACGTATCTTCTAAAAATTACTCATGCCA 2282
Db   7865 CTCCAAAACTAGAGAACGTCAATCTCCAAGTACGACTTCCTCCCACAACTACTTTGCAGG 7806

Qy   2283 AGGAGAAATGCTCTTCTCATTGCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTTTA 2342
Db   7805 ACTCCGCTTCGACAGTCTCCTCTTCAGGGACTTCATCTCTACAGGGCTATCCCTAGGTTA 7746

Qy   2343 CAGCTATGGAGACCATAACTGTCACCATTTCTATACCCAAGGAGAAAATCTAACATCTCA 2402
Db   7745 TAGCTACGGAGATCACCATATGCTTTGCCACTATACAGAAATCTTAAAAGGGTCGTCCAA 7686

Qy   2403 AGGGACGTTCCGTAGTCAAACGATGGGAGGTGCTGTTTTTTTGATCTCCCTATGAAACC 2462
Db   7685 AGCCTTCTTTAATAACCACACTTTGGTAGCCTCTCTAGACTGCACATTCTTACCAGCTAG 7626

Qy   2463 CTTTGGATCAACGCATATACTGACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCTAG 2522
Db   7625 AATCACCCGCACTCTCGAACTCCA---GCCCTTTATCAGTGCCATTGCTCTGCGCTGTTC 7569

Qy   2523 CCTGTCTCACTTTACTGAGGTGGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCTTT 2582
Db   7568 CCAGGCCTCGTTCCAAGAAACTGGAGACCATATAAGAAAATTCCATCCAAAACATCCCCT 7509

Qy   2583 GATCAATGTCCTAGTCCCTATTGGAGTTAAAGGTAGCTTTATGAATGCTACCCAAAGACC 2642
Db   7508 TACAGATCTTTCCTCTCCCATAGGCTTCCGTTCTGAATGGAAAACTTCACATCATATCCC 7449

Qy   2643 TCAAGCCTGGACTGTAGAATTGGCATACCAACCCGTTCTGTATAGACAAGAACCAGGGAT 2702
Db   7448 CATGCTATGGACTACGGAAATATCCTACGTACCTACCCTATACAGAAAAAATCCAGAAAT 7389

Qy   2703 CGCGACCCAGCTCCTAGCCAGTAAGGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCGCG 2762
Db   7388 GTTCACGACACTACTCATCAGCAATGGAACATGGACAACACAAGCAACTCCCGTCTCCTA 7329

Qy   2763 TCATGCCATGTCCTATAAAATCTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTCCA 2822
Db   7328 TAATTCCGTAGCTGCAAAAATAAAAAATACTTCCCAACTTTTCTCAAGAGTAACCTTATC 7269

Qy   2823 TTTCCAGTATCATGGATTCTACTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAAAT 2882
Db   7268 CTTAGATTATTCAGCTCAAGTCTCCTCGTCAACTGTAGGTCAATACCTTAAAGCTGAGAG 7209

Qy   2883 TGCTCTGCGATTCTA 2897
Db   7208 TCATTGCACATTTTA 7194
```

```
RESULT 13
AX349483
LOCUS       AX349483                2817 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 6 from Patent WO0202606.
ACCESSION   AX349483
VERSION     AX349483.1  GI:18615348
KEYWORDS    .
SOURCE      Chlamydophila pneumoniae.
  ORGANISM  Chlamydophila pneumoniae
            Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
REFERENCE   1
  AUTHORS   Ratti,G. and Grandi,G.
  TITLE     Immunisation against Chlamydia pneumoniae
  JOURNAL   Patent: WO 0202606-A 6 10-JAN-2002;
            Chiron S.p.A. (IT)
FEATURES             Location/Qualifiers
     source          1..2817
                     /organism="Chlamydophila pneumoniae"
                     /db_xref="taxon:83558"
BASE COUNT      800 a    706 c    534 g    777 t
ORIGIN

  Query Match             3.0%;  Score 88;  DB 6;  Length 2817;
  Best Local Similarity   43.4%;  Pred. No. 1.2e-09;
  Matches  882;  Conservative    0;  Mismatches 1095;  Indels   57;  Gaps    8;

Qy    882 CAATCGTGGTAGGATCTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTA 941
Db    822 CAACCGAGAAGGGGTACTCTTTAACAATAACCAAAGCCAGAGCAACGGTGGAGCCATTCA 881

Qy    942 CGCTCCTGTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAA 1001
Db    882 TGCGAAATCTATCATTATCAAAGAAAATGGTCCTGTATACTTTTAAATAACACTGCAAC 941

Qy   1002 TAATAAGGGGGGCGCTATCTATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCG 1061
Db    942 TCGGGGAGGGGCTCTCCTCAACTTA--------------TCAGCAGGTTCTGGAAACGG 986

Qy   1062 CCATGCTATTATTTTTAATGAAAATATTGTGACTAATGTAACTAATGCAAATGGTACCAG 1121
Db    987 AAGCTTCATCTTATCTGCAGATAATGGAGATATTATCTTTAACAATAATACGGCCTCCAA 1046

Qy   1122 TACGTCAGCTAATCCTCCTAGAAGAAATGCAATAACAGTAGCAAGCTCCTCTGGTGAAAT 1181
Db   1047 GCATGCCCTCAATCCTCCATACAGAAACGCCATTCAC---TCGACTCCTAATATGAATCT 1103

Qy   1182 TCTATTAGGAGCAGGGAGTAGCCAAAATTTAATTTTTTATGATCCTATTGAAGTT---AG 1238
Db   1104 GCAAATAGGAGCCCGTCCCGGCTATCGAGTGCTGTTCTATGATCCCATAGAACATGAGCT 1163

Qy   1239 CAATGCAGGGGTCTCTGTGTCCTTCAATAAGGAAGCTGATCAAACAGGCTCTGTAGTATT 1298
Db   1164 CCCTTCCTCCTTCCCCATACTCTTTAATTTCGAAACCGGTCATACAGGTACAGTTTTATT 1223
```

```
Qy   1299 TTCAGGAGCTACTGTTAATTCTGCAGATTTTCATCAACGCAATTTACAAACAAAAACACC 1358
Db   1224 TTCAGGGGAACATGTACACCAGAACTTTACCGATGAAATGAATTTCTTTTCCTATTTAAG 1283

Qy   1359 TGCACCCCTTACTCTCAGTAATGGTTTTCTATGTATCGAAGATCATGCTCAGCTTACAGT 1418
Db   1284 GAACACTTCGGAACTACGTCAAGGAGTCCTTGCTGTTGAAGATGGTGCGGGGCTGGCCTG 1343

Qy   1419 GAATCGATTCACACAAACTGGGGGTGTTGTTTCTCTTGGGAATGGAGCAGTTCTGAGTTG 1478
Db   1344 CTATAAGTTCTTCCAACGAGGAGGCACTCTACTTCTAGGTCAAGGTGCGGTGATCACGAC 1403

Qy   1479 CTATAAAAATGGTGCAGGAAATTCTGCTAGCAATGC----------CTCTATAACACT 1526
Db   1404 AGCAGGAACGATTCCCACACCATCCTCAACACCAACGACAGTAGGAAGTACTATAACTTT 1463

Qy   1527 GAAGCATATTGGATTGAATCTTTCTTCCATTCTGAAAAGTGGTGCTGAGATTCCTTTATT 1586
Db   1464 AAATCACATTGCCATTGACCTTCCTTCTATTCTTTCTTTTCAAGCTCAGGCTCCAAAAAT 1523

Qy   1587 GTGGGTAGAGCCTACAAATAACAGCAATAACTATACAGCAGATACTGCAGCTACCTTTTC 1646
Db   1524 TTGGATTTACCCCACAAAAACAGGATCTACCTATACTGAAGATTCCAACCCGACAATCAC 1583

Qy   1647 ATTAAGTGATGTAAAACTCTCACTCATTGATGACTATGGGAATTCTCCTTATGAATCCAC 1706
Db   1584 ---AATCTCAGGAACTCTCACCTTACGCAACAGCAACAACGAAGATCCCTACGATAGTCT 1640

Qy   1707 AGATCTAACCCATGCTCTGTCATCACAGCCTATGCTATCTATTTCTGAGGCTAGTGATAA 1766
Db   1641 GGATCTCTCGCACTCTCTTGAGAAAGTTCCCCTTCTTTATATTGTCGATGTCGCTGCACA 1700

Qy   1767 CCAGCTAAGATCTGATGATATGGATTTTTCGGGACTAAA---TGTCCCTCATTATGGATG 1823
Db   1701 AAAAATTAACTCTTCGCAACTGGATCTATCCACATTAAATTCTGGCGAACACTATGGGTA 1760

Qy   1824 GCAAGGACTTTGGACTTGGGGCTGGGCAAAAACTCAAGATCCAGAACCAGCATCTTCAGC 1883
Db   1761 TCAAGGCATCTGGTCGACCTATTGGGTAGAAACTA--------------CAACAATCAC 1805

Qy   1884 AACAATCACAGATCCACAAAAAGCCAATAGATTCCATAGAACCTTATTACTGACTTGGCT 1943
Db   1806 GAACCCTACATCTCTACTAGGCGCGAATACAAAACACAAGCTGCTCTATGCAAACTGGTC 1865

Qy   1944 TCCTGCTGGGTATGTTCCTAGCCCGAAACACAGAAGTCCCCTCATAGCGAATACCTTATG 2003
Db   1866 TCCTCTAGGCTACCGTCCTCATCCCGAACGTCGAGGAGAATTCATTACGAATGCCTTGTG 1925

Qy   2004 GGGGAATATGCTGCTTGCAACAGAAAGCTTAAAAAATAGTGCAGAACTGACACCTAGTGA 2063
Db   1926 GCAATCGGCATATACGGCTCTTGCAGGACTCCACTCCCTCTCCTCCTGGGATGAAGAGAA 1985

Qy   2064 TCATCCTTTCTGGGGAATTACAGGAGGAGGACTAGGCATGATGGTTTACCAAGATCCTCG 2123
Db   1986 GGGTCATGCAGCTTCCCTACAAGGCATTGGTCTTCTGGTTCATCAAAAAGACAAAAACGG 2045

Qy   2124 AGAAAATCATCCTGGATTCCATATGCGCTCTTCCGGATACTCTGCGGGGATGATAGCAGG 2183
Db   2046 TTTTAAGGGATTTCGTAGTCATATGACAGGTTATAGTGCTACCACCGAAGCAACCTCTTC 2105

Qy   2184 GCAGACACACACCTTCTCATTGAAATTCAGTCAGACCTACACCAAACTCAATGAGCGTTA 2243
Db   2106 TCAAAGTCCGAATTTCTCTTTAGGATTTGCTCAGTTCTTCTCCAAAGCTAAAGAACATGA 2165

Qy   2244 CGCAAAAAACAACGTATCTTCTAAAAATTACTCATGCCAAGGAGAAATGCTCTTCTCATT 2303
Db   2166 ATCTCAAAATAGCACGTCCTCTCACCACTATTTCTCTGGAATGTGCATAGAAAATACTCT 2225

Qy   2304 GCAAGAAGGTTTCTTGCTGACTAAATTAGTTGGGCTTTACAGCTATGGAGACCATAACTG 2363
Db   2226 CTTCAAAGAGTGGATACGTCTATCTGTGTCTCTTGCTTATATGTTTACCTCGGAACATAC 2285

Qy   2364 TCACCATTTCTATACCCAAGGAGAAAATCTAACATCTCAAGGGACGTTCCGTAGTCAAAC 2423
Db   2286 CCATACAATGTATCAGGGTCTCCTGGAAGGGAACTCTCAGGGATCTTTCCACAACCATAC 2345

Qy   2424 GATGGGAGGTGCTGTTTTTTTTGATCTCCCTATGAAACCCTTTGGATCAACGCATATACT 2483
Db   2346 CTTAGCAGGGGCTCTCTCCTGTGTTTTCTTACCTCAACCTCACGGCGAGTCCC---TGCA 2402

Qy   2484 GACAGCTCCCTTTTTAGGTGCTCTTGGTATTTATTCTAGCCTGTCTCACTTTACTGAGGT 2543
Db   2403 GATCTATCCCTTTATTACTGCCTTAGCCATCCGAGGAAATCTTGCTGCGTTTCAAGAATC 2462

Qy   2544 GGGAGCCTATCCGCGAAGCTTTTCTACAAAGACTCCTTTGATCAATGTCCTAGTCCCTAT 2603
Db   2463 TGGAGACCATGCTCGGGAATTTTCCCTACACCGCCCCCTAACGGACGTCTCCCTCCCTGT 2522

Qy   2604 TGGAGTTAAAGGTAGCTTTATGAATGCTACCCAAAGACCTCAAGCCTGGACTGTAGAATT 2663
Db   2523 AGGAATCCGCGCTTCTTGGAAGAACCACCACCGAGTTCCCCTAGTCTGGCTCACAGAAAT 2582

Qy   2664 GGCATACCAACCCGTTCTGTATAGACAAGAACCAGGGATCGCGACCCAGCTCCTAGCCAG 2723
Db   2583 TTCCTATCGCTCTACTCTCTATAGGCAAGATCCTGAACTCCACTCGAAATTACTGATTAG 2642

Qy   2724 TAAGGGTATTTGGTTTGGTAGTGGAAGCCCCTCATCGCGTCATGCCATGTCCTATAAAAT 2783
Db   2643 CCAAGGTACGTGGACGACGCAGGCCACTCCTGTGACCTACAATGCTTTAGGGATCAAAGT 2702

Qy   2784 CTCACAGCAAACACAACCTTTGAGTTGGTTAACTCTCCATTTCCAGTATCATGGATTCTA 2843
Db   2703 GAAAAATACCATGCAGGTGTTTCCTAAAGTCACTCTCTCCTTAGATTACTCTGCGGATAT 2762

Qy   2844 CTCCTCTTCAACCTTCTGTAATTATCTCAATGGGGAAATTGCTCTGCGATTCTA 2897
Db   2763 TTCTTCCTCCACGCTGAGTCACTACTTAAACGTGGCGAGTAGAATGAGATTTTA 2816
```

```
RESULT 14
AE002293
LOCUS       AE002293               14482 bp    DNA     linear   BCT 26-MAY-2000
DEFINITION  Chlamydia muridarum, section 25 of 85 of the complete genome.
ACCESSION   AE002293 AE002160
VERSION     AE002293.1  GI:7190298
KEYWORDS    .
SOURCE      Chlamydia muridarum.
  ORGANISM  Chlamydia muridarum
            Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE   1  (bases 1 to 14482)
  AUTHORS   Read,T.D., Brunham,R., Shen,C., Gill,S.R., Heidelberg,J.F.,
            White,O., Hickey,E.K., Peterson,J., Umayam,L.A., Utterback,T.,
            Berry,K., Bass,S., Linher,K., Weidman,J., Khouri,H., Craven,B.,
            Bowman,C., Dodson,R., Gwinn,M., Nelson,W., DeBoy,R., Kolonay,J.,
            McClarty,G., Salzberg,S.L., Eisen,J. and Fraser,C.M.
  TITLE     Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
            pneumoniae AR39
  JOURNAL   Nucleic Acids Res. 28 (6), 1397-1406 (2000)
  MEDLINE   20150255
   PUBMED   10684935
REFERENCE   2  (bases 1 to 14482)
  AUTHORS   Read,T.D., Brunham,R., Shen,C., Gill,S.R., Heidelberg,J.F.,
            White,O., Hickey,E.K., Peterson,J., Umayam,L.A., Utterback,T.,
            Berry,K., Bass,S., Linher,K., Weidman,J., Khouri,H., Craven,B.,
            Bowman,C., Dodson,R., Gwinn,M., Nelson,W., DeBoy,R., Kolonay,J.,
            McClarty,G., Salzberg,S.L., Eisen,J. and Fraser,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAR-2000) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .14482
                     /organism="Chlamydia muridarum"
                     /db_xref="taxon:83560"
                     /note="synonym: Chlamydia trachomatis MoPn"
     gene            complement(355. .825)
```

/gene="TC0259"
CDS complement(355. .825)
/gene="TC0259"
/note="hypothetical protein; identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF39128.1"
/db_xref="GI:7190299"
/translation="MVSPSYPSAKTISASLHESPQQLLVQKRNRWSFLQKVVRIALAV
LVVIMTLGLILFFYTFSEITTFPWCCRKHLLPKEDIIYPLASPSPSLQHLPSIHEEKA
PAPKTPELLKAKPYISTYPLLESINTKESLVPIISEKVVLEFFEEISTRKRIQS"
gene complement(937. .2046)
/gene="TC0260"
CDS complement(937. .2046)
/gene="TC0260"
/note="hypothetical protein; identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF39129.1"
/db_xref="GI:7190300"
/translation="MAKISSFSSLPTFDAPKSPPSLFSQSSLLGKIARIIVALFLILV
SLGLILIAYRFSDLLNCKFCITKTTELPIPSTKIPISSPPVPTTPPYQKKEPTLPQKN
PKISENLLSQEVVQDYLSSGRLPELAILDNSQMFQFMCVLHDQYPKLLPNDCLIPLTI
FNYREEICNTIQDKLKADQGQYCSLGDLQCPITCSPENYHQLLQQSRVLPFLLWYDPE
PTNHQQTLEKMQEIASQGTPGNSHWTVIVVDLDARCITYFDSLVNYIASTDEMERRMK
SLACCLANIGLCKNNGCPFDVHVAVNESLQNWMGSCCGLWCCQYMKWYMDHSHTGILQ
KIPDSLAYKTLLLQSLHSTFEKLMKKYADLSWPTT"
gene complement(2125. .5055)
/gene="TC0261"
CDS complement(2125. .5055)
/gene="TC0261"
/note="similar to GP:4376753; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="polymorphic membrane protein E/F family"
/protein_id="AAF39130.1"
/db_xref="GI:7190301"
/translation="MKKLFFFVLIGSSILGFTREVPPSILLKPILNPYHMTGLFFPKV
NLLGDTHNLTDYHLDNLKCILACLQRTPYEGAAFTVTDYLGFSDTQKDGIFCFKNLTP
ESGGVIGSPTQNTPTIKIHNTIGPVLFENNTCHRLWTQTDPENEGNKAREGGAIHAGD
VYISNNQNLVGFIKNFAYVQGGAISANTFAYKENKSSFLCLNNSCIQTKTGGKGGAIY
VSTSCSFENNNKDLLFIQNSGCAGGAIFSPTCSLIGNQGDIVFYSNHGFKNVDNATNE
SGDGGAIKVTTRLDITNNGSQIFFSDNISRNFGGAIHAPCLHLVGNGPTYFTNNIANH
TGGAIYITGTETSKISADHHAIIFDNNISANATNADGSSSNTNPPHRNAITMDNSAGG
IELGAGKSQNLIFYDPIQVTNAGVTVDFNKDASQTGCVVFSGATVLSADISQANLQTK
TPATLTLSHGLLCIEDRAQLTVNNFTQTGGIVALGNGAVLSSYQHSTTDATQTPPTTT
TTDASVTLNHIGLNLPSILKDGAEMPLLWVEPISTTQGNTTTYTSDTAASFSLNGATL
SLIDEDGNSPYENTDLSRALYAQPMLAISEASDNQLQSESMDFSKVNVPHYGWQGLWT
WGWAKTENPTTTPPATITDPKKANQFHRTLLLTWLPAGYIPSPKHKSPLIANTLWGNI
LFATENLKNSSGQELLDRPFWGITGGGLGMMVYQEPRKDHPGFHMHTSGYSAGMITGN
THTFSLRFSQSYTKLNERYAKNYVSSKNYSCQGEMLLSLQEGLMLTKLIGLYSYGNHN
SHHFYTQGEDLSSQGEFHSQTFGGAVFFDLPLKPFGRTHILTAPFLGAIGMYSKLSSF
TEVGAYPRTFITETPLINVLIPIGVKGSFMNATHRPQAWTVELAYQPVLYRQEPSIST
QLLAGKGMWFGHGSPASRHALAYKISQKTQLLRFATLQLQYHGYYSSSTFCNYLNGEV
SLRF"
gene complement(5058. .8135)
/gene="TC0262"
CDS complement(5058. .8135)
/gene="TC0262"
/note="similar to GP:4376753; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="polymorphic membrane protein E/F family"
/protein_id="AAF39131.1"
/db_xref="GI:7190302"
/translation="MTRRILPLSLVFIPLSCISASETDTLKLPNLTFGGREIEFIVTP
PSSIAAQYITYANVSNYRGNFTISSCTQDQWFSRGLSTTNSSGAFVESMTSFTAIDNA
DLFFCNNYCTHQGGGGAINATGLISFKNNQNILFYNNTTIGTQFTGVALRTERNRGGA
LYGSSIELINNHSLNFINNTSGDMGGAVSTIQNLVIKNTSGIVAFENNHTTDHIPNTF
ATILARGGAVGCQGACEISHNTGPVVFNSNYGGYGGAISTGGQCIFRDNKDKLIFINN
SALGWHNTSAQGNGAVISAGGEFGLLNNKGPIYFENNNASYIAGAISCNNLNFQENGP
IYFLNNSALYGGAFHLFASPAANYIHTGSGDIIFNNNTELSTTGMSAGLRKLFYIPGT
TNNNPITLSLGAKKDTRIYFYDLFQWGGLKKANTPPENSPHTVTINPSDEFSGAVVFS
YKNISSDLQAHMIASKTHNQIKDSPTTLKFGTMSIENGAEFEFFNGPLTQESTSLLAL
GQDSILTVGKDASLTITHLGIILPGLLNDQGTTAPRIRVNPQDMTQNTNSNQAPVSTE
NVATQKIFFSGLVSLVDENYESVYDSCDLSRGKANQPILHIETTNDAQLSNDWKNTLN
TSLYSLPHYGYQGLWTSNWMTTTRTVSLTNSTETQTANNSIQEQKNTSETFDSNSTTT
AKIPSIRASTGGTTPLATTDVTVTRHSLVVSWTPIGYIADPARRGDLIANNLVSSGRN
TTLYLRSLLPDDSWFALQGSAATLFTKQQKRLDYHGYSSASKGYAISSQASGAHGHKF
LFSFSQSSDTMKEKRTNNKISSRYYLSALCFEQPMFDRIALIGAAAYNYGTHKTYNFY
GTKKFSKGNFHSTTLGGSLRCELRDSMPFQSIMLTPFIQALISRTEPASIQEQGDLAR
LFSLKQPHTAVVSPIGIKGVYSSNKWPTVSCEMEVAYQPTLYWKRPILNTVLIKNNGS
WETTNTPLAKHSFYGRGSSSLKFSYLKLFANYQAQVATSTVSHYMNAGGALVF"
gene 8307. .11270
/gene="TC0263"
CDS 8307. .11270
/gene="TC0263"
/note="similar to GP:4376736; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="polymorphic membrane protein G family"
/protein_id="AAF39132.1"
/db_xref="GI:7190303"
/translation="MMQTPFHKFFLLAMLSYSLLQGGHAADISMPPGIYDGTTLTAPF
PYTVIGDPRGTKVTSSGSLELKNLDNSIATLPLSCFGNLLGNFTIAGRGHSLVFENIR
TSTNGAALSNHAPSGLFVIEAFDELSLLNCNSLVSVVPQTGGTTTSVPSNGTIYSRTD
LVLRDIKKVSFYSNLVSGDGGAIDAQSLMVNGIEKLCTFQENVAQSDGGACQVTKTFS
AVGNKVPLSFLGNVAGNKGGGVAAVKDGQGAGGATDLSVNFANNTAVEFEGNSARIGG
GIYSDGNISFLGNAKTVFLSNVASPIYVDPAAAGGQPPADKDNYGDGGAIFCKNDTNI
GEVSFKDEGVVFFSKNIAAGKGGAIYAKKLTISDCGPVQFLGNVANDGGAIYLVDQGE
LSLSADRGDIIFDGNLKRMATQGAATVHDVMVASNAISMATGGQITTLRAKEGRRILF
NDPIEMANGQPVIQTLTVNEGEGYTGDIVFAKGDNVLYSSIELSQGRIILREQTKLLV
NSLTQTGGSVHMEGGSTLDFAVTTPPAANSMALTNVHFSLASLLKNNGVTNPPTNPPV
QVSSPAVIGNTAAGTVTISGPIFFEDLDETAYDNNQWLGADQTIDVLQLHLGANPPAN
APTDLTLGNESSKYGYQGSWTLQWEPDPANPPQNNSYMLKASWTKTGYNPGPERVASL
VSNSLWGSILDVRSAHSAIQASIDGRAYCRGIWISGISNFFYHDQDALGQGYRHISGG
YSIGANSYFGSSMFGLAFTETFGRSKDYVVCRSNDHTCVGSVYLSTRQALCGSCLFGD
AFVRASYGFGNQHMKTSYTFAEESNVRWDNNCVVGEVGAGLPIMLAASKLYLNELRPF
VQAEFAYAEHESFTERGDQAREFKSGHLMNLSIPVGVKFDRCSSKHPNKYSFMGAYIC
DAYRSISGTETTLLSHKETWTTDAFHLARHGVMVRGSMYASLTGNIEVYGHGKYEYRD
ASRGYGLSIGSKIRF"
gene 11292. .14243
/gene="TC0264"
CDS 11292. .14243
/gene="TC0264"
/note="similar to GP:4376737; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="polymorphic membrane protein H family"
/protein_id="AAF39133.1"
/db_xref="GI:7190304"
/translation="MLVMPFSLRSTSFCFLACLCSYSYGLASSPQVLTPNVIIPFKGD
DIYLNGDCVFASIYAGAEQGSIISANGQNLTIVGQNHTLSFTDSQGPALQNCAFISAE
EKISLRDFSSLLFSKNVSCGEKGMISGKTVSISGGDSIVFKDNSVGYSSLPSVGQTPT
TPIVGDVLKGSIFCVETGLEISGVKKELVFDNTAGNFGAVFCSRAAQGDTTFTVKDCK
GKILFQDNVGSCGGGVIYKGEVLFQDNEGEMLFRGNSAHDDLGILDANPQPPTEVGGG
GGVICTPEKTVTFKGNKGPITFDYNFAKGRGGAIQSQTFSLVADSAVVFSNNTAEKGG
GAIYALEVNVSTNGGSILFEGNRASEGGAICVSEPIAANNGGLTLHAADGDIIFSKNM
TSDRPGERSAIRILDSGTNVSLNASGASKMIFYDPVVQNNPATPPTGTSGEIKINESG
SGSVVFTAETLTPSEKLNVINATSNFPGNLTVSSGELVVTKGATLTVGNITATSGRVT
LGSGASLSAVAGTAGTCTVSKLGIDLESFLVPTYETAKLGADTTVAVNNNPTLDLVMA
NETEMYDNPLFMNAVTIPFVTLVSLQTTGGVTTSAVTLNNADTAHYGYQGSWSADWRR
PPLAPDPSGMTPLDKSNTLYVTWRPSSNYGVYKLDPQRRGELVPNSLWVSGSALRTFT
NGLKEHYVSRDVGFIASVQALGDYVLNYKQGNRDGFLARYGGFQAVAASHYENGGIFG
VAFGQLYGQTKSRLYDSKDAGNITILSCFGRSYIDVKGTETVVYWETAYGYSVHRMHT
QYFNGKTNKFDHSKCRWHNNSYYAFVGAEHNFLEYCIPTRQLARDYDLTGFMRFEMSG
GWSSGAKETGALPRHFDRGTGHNMSLPIGVVAHAVSNGRRSPPSKLTINMGYRPDIWR
VTPHCNMKIIANGVKTPIQGSPLARHAFFLEVHDTLYVRHLGRAYMNYSLDARHRQTT

```
                    HFVSLGLNRIF"
BASE COUNT     4185 a   2517 c   3329 g   4451 t
ORIGIN

  Query Match             2.0%;  Score 56.6;  DB 1;  Length 14482;
  Best Local Similarity   55.8%;  Pred. No. 0.018;
  Matches  129;  Conservative    0;  Mismatches   99;  Indels    3;  Gaps    1;

Qy    889 GGTAGGATCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCT 948
          |||    | || |||||| | |||||  |   |    | || || |||||||| |||
Db   9333 GGTGTTGTTTTCTTTAGTAAAAATATTGCCGCAGGAAAGGGGGGCGCTATTTATGCTAAG 9392

Qy    949 GTAGTTACCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAG 1008
          | | ||  |    ||   ||| ||   | | ||| |    ||| |||| ||| |
Db   9393 AAACTGACAATTTCTGACTGTGGTCCGGTCCAGTTTCTTGGTAATGTCGCGAATGAC--- 9449

Qy   1009 GGGGGCGCTATCTATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCT 1068
          |||||||||| ||| |||  |      |  |    |   | ||||| || |||   | |
Db   9450 GGGGGCGCTATTTATCTAGTAGATCAGGGGGAACTTAGTCTATCTGCTGATCGCGGAGAT 9509

Qy   1069 ATTATTTTTAATGAAAATATTGTGACTAATGTAACTAATGCAAATGGTACC 1119
          ||||||||| ||| |||| |   ||  |  |  ||  | |    ||  |||
Db   9510 ATTATTTTTGATGGAAATTTAAAGAGAATGGCTACGCAAGGCGCTGCCACC 9560


RESULT 15
AX155927
LOCUS       AX155927               2949 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 170 from Patent WO0140474.
ACCESSION   AX155927
VERSION     AX155927.1  GI:14537032
KEYWORDS    .
SOURCE      Chlamydia sp.
  ORGANISM  Chlamydia sp.
            Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE   1  (bases 1 to 2949)
  AUTHORS   Probst,P., Bhatia,A., Skeiky,Y.A., Fling,S.P. and Scholler,J.
  TITLE     Compounds and methods for treatment and diagnosis of chlamydial
            infection
  JOURNAL   Patent: WO 0140474-A 170 07-JUN-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .2949
                     /organism="Chlamydia sp."
                     /db_xref="taxon:35827"
BASE COUNT      811 a    540 c    704 g    892 t      2 others
ORIGIN

  Query Match             1.7%;  Score 48.4;  DB 6;  Length 2949;
  Best Local Similarity   56.7%;  Pred. No. 1.7;
  Matches  110;  Conservative    0;  Mismatches   81;  Indels    3;  Gaps    1;

Qy    896 TCTTTTTTAGTGACAATATCACAAAAAATTATGGCGGAGCTATTTACGCTCCTGTAGTTA 955
          | || ||||||  |||| |  |         | || ||||||||||| ||        |
Db   1001 TTTTCTTTAGTAGCAATGTAGCTGCTGGGAAAGGGGGAGCTATTTATGCCAAAAAGCTCT 1060

Qy    956 CCCTAGTGGATAATGGCCCTACCTACTTTATAAACAATATCGCCAATAATAAGGGGGGCG 1015
          |  | |   |  |||||||    | ||| |||  |||||||| ||| ||   || || |
Db   1061 CGGTTGCTAACTGTGGCCCTGTACAATTTTTAAGGAATATCGCTAATGAT---GGTGGAG 1117

Qy   1016 CTATCTATATAGACGGAACCAGCAACTCCAAAATTTCTGCCGACCGCCATGCTATTATTT 1075
          | || ||| |||  | | |  |  |   ||   | ||||| ||       | ||||||||
Db   1118 CGATTTATTTAGGAGAATCTGGAGAGCTCAGTTTATCTGCTGATTATGGAGATATTATTT 1177

Qy   1076 TTAATGAAAATATT 1089
          |  |||  ||| ||
Db   1178 TCGATGGGAATCTT 1191


Search completed: February  4, 2003, 18:21:27
Job time : 8000 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2003, 20:59:53 ; Search time 75 Seconds
(without alignments)
285.245 Million cell updates/sec

Title: US-09-677-752-2
Perfect score: 5086
Sequence: 1 MKKAFFFFLIGNSLSGLARE.....GFYSSSTFCNYLNGEIALRF 965

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 129505 seqs, 22169297 residues

Total number of hits satisfying chosen parameters: 129505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4994.5	98.2	964	10	US-09-841-132-177	Sequence 177, App
2	4919.5	96.7	977	10	US-09-841-132-191	Sequence 191, App
3	989.5	19.5	963	10	US-09-886-468-22	Sequence 22, Appl
4	647	12.7	982	10	US-09-841-132-176	Sequence 176, App
5	647	12.7	1006	10	US-09-841-132-190	Sequence 190, App
6	645	12.7	925	10	US-09-452-380-4	Sequence 4, Appli
7	645	12.7	936	10	US-09-452-380-3	Sequence 3, Appli
8	636	12.5	926	9	US-10-023-437-57	Sequence 57, Appl
9	605.5	11.9	715	10	US-09-841-132-321	Sequence 321, App
10	595	11.7	1016	12	US-10-007-693-95	Sequence 95, Appl
11	583	11.5	978	12	US-10-007-693-65	Sequence 65, Appl
12	569.5	11.2	922	10	US-09-886-468-19	Sequence 19, Appl
13	531	10.4	839	9	US-10-023-437-23	Sequence 23, Appl
14	517	10.2	1723	10	US-09-841-132-394	Sequence 394, App
15	517	10.2	1723	10	US-09-841-132-395	Sequence 395, App
16	474.5	9.3	1609	12	US-10-007-693-75	Sequence 75, Appl
17	471.5	9.3	880	10	US-09-841-132-175	Sequence 175, App
18	459.5	9.0	866	10	US-09-841-132-189	Sequence 189, App
19	456.5	9.0	931	10	US-09-779-081-2	Sequence 2, Appli
20	398.5	7.8	1752	10	US-09-841-132-180	Sequence 180, App
21	395	7.8	1751	10	US-09-841-132-445	Sequence 445, App
22	395	7.8	1751	10	US-09-841-132-594	Sequence 594, App
23	388.5	7.6	679	9	US-10-023-437-53	Sequence 53, Appl
24	386.5	7.6	1530	10	US-09-841-132-178	Sequence 178, App
25	386	7.6	1531	12	US-10-007-693-98	Sequence 98, Appl
26	371.5	7.3	715	10	US-09-841-132-329	Sequence 329, App
27	365	7.2	848	10	US-09-841-132-192	Sequence 192, App
28	363	7.1	1776	10	US-09-841-132-179	Sequence 179, App
29	345.5	6.8	948	10	US-09-841-132-194	Sequence 194, App
30	342	6.7	1770	10	US-09-841-132-444	Sequence 444, App
31	340	6.7	646	10	US-09-841-132-317	Sequence 317, App
32	320.5	6.3	691	10	US-09-841-132-313	Sequence 313, App
33	278.5	5.5	631	10	US-09-841-132-325	Sequence 325, App
34	278	5.5	871	10	US-09-886-468-21	Sequence 21, Appl
35	277.5	5.5	439	10	US-09-841-132-524	Sequence 524, App
36	246.5	4.8	514	10	US-09-886-468-23	Sequence 23, Appl
37	246.5	4.8	619	10	US-09-841-132-309	Sequence 309, App
38	235	4.6	700	10	US-09-841-132-345	Sequence 345, App
39	234	4.6	375	10	US-09-886-468-20	Sequence 20, Appl
40	221.5	4.4	821	10	US-09-841-132-195	Sequence 195, App
41	220.5	4.3	654	10	US-09-841-132-341	Sequence 341, App
42	215	4.2	778	10	US-09-841-132-193	Sequence 193, App
43	213	4.2	469	10	US-09-886-468-18	Sequence 18, Appl
44	205	4.0	683	10	US-09-841-132-357	Sequence 357, App
45	188.5	3.7	583	10	US-09-841-132-353	Sequence 353, App

ALIGNMENTS

RESULT 1
US-09-841-132-177
; Sequence 177, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 177
; LENGTH: 964
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-841-132-177

Query Match 98.2%; Score 4994.5; DB 10; Length 964;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 950; Conservative 7; Mismatches 7; Indels 1; Gaps 1;

```
Qy       1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 60

Qy      61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120

Qy     121 RDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVR 180
           |||||||||||||||| || || || :|||||||||||||||||||||||||||||||||:
Db     121 RDTIGPVIFENNTCCRLFTWRNPYAA-DKIREGGAIHAQNLYINHNHDVVGFMKNFSYVQ 179

Qy     181 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 239

Qy     241 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 300
```

```
Db    240 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 299

Qy    301 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFN 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFN 359

Qy    361 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 419

Qy    421 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 479

Qy    481 GVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNY 540
          |||||||||||||||| |:|||||||||||||||||||||||||||||||||||||||||||
Db    480 GVVSLGNGAVLSCYKNGTGDSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNY 539

Qy    541 TADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||:|:::|||||||||
Db    540 TADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLQSENIDFSG 599

Qy    601 LNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRS 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 LNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRS 659

Qy    661 PLIANTLWGNMLLATESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSG 720
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db    660 PLIANTLWGNMLLATESLKNSAELTPSGHPFWGITGGGLGMMVYQDPRENHPGFHMRSSG 719

Qy    721 YSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGL 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    720 YSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGL 779

Qy    781 YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYS 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    780 YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYS 839

Qy    841 SLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPG 900
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    840 SLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATHRPQAWTVELAYQPVLYRQEPG 899

Qy    901 IATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGE 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    900 IATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGE 959

Qy    961 IALRF 965
          |||||
Db    960 IALRF 964
```

RESULT 2
US-09-841-132-191
; Sequence 191, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 191
; LENGTH: 977
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-841-132-191

```
  Query Match           96.7%;  Score 4919.5;  DB 10;  Length 977;
  Best Local Similarity 98.4%;  Pred. No. 0;
  Matches  935;  Conservative    7;  Mismatches    7;  Indels    1;  Gaps    1;

Qy     16 GLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEG 75
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     29 GLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEG 88

Qy     76 AAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCC 135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     89 AAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCC 148

Qy    136 RPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSEN 195
          | ||  || || :|||||||||||||||||||||||||||||:|||||||||||||||||
Db    149 RLFTWRNPYAA-DKIREGGAIHAQNLYINHNHDVVGFMKNFSYVQGGAISTANTFVVSEN 207

Qy    196 QSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNACCAGGAIFSPICSLTG 255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    208 QSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNACCAGGAIFSPICSLTG 267

Qy    256 NRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYA 315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    268 NRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYA 327

Qy    316 PVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTST 375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    328 PVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTST 387

Qy    376 SANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVSFNKEADQTGSVVFSG 435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    388 SANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVSFNKEADQTGSVVFSG 447

Qy    436 ATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYK 495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    448 ATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYK 507

Qy    496 NGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKL 555
          || |:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    508 NGTGDSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKL 567

Qy    556 SLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSGLNVPHYGWQGLWTWG 615
          ||||||||||||||||||||||||||||||||||:|:::|||||||||||||||||||||
Db    568 SLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLQSENIDFSGLNVPHYGWQGLWTWG 627

Qy    616 WAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLAT 675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    628 WAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLAT 687

Qy    676 ESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFSL 735
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    688 ESLKNSAELTPSGHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFSL 747

Qy    736 KFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQG 795
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    748 KFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQG 807

Qy    796 ENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSF 855
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    808 ENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSF 867

Qy    856 STKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGS 915
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    868 STKTPLINVLVPIGVKGSFMNATHRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGS 927

Qy    916 GSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    928 GSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 977
```

RESULT 3
US-09-886-468-22
; Sequence 22, Application US/09886468
; Patent No. US20020037293A1

```
; GENERAL INFORMATION:
;  APPLICANT: Aventis Pasteur Limited
;  TITLE OF INVENTION: Chlamydia antigens and corresponding DNA fragments and uses there
;  FILE REFERENCE: 77813-5
;  CURRENT APPLICATION NUMBER: US/09/886,468
;  CURRENT FILING DATE:  1999-12-23
;  PRIOR APPLICATION NUMBER: 60/113,280
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: 60/113,281
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: 60/113,282
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: 60/113,283
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: 60/113,284
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: 60/113,285
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: 60/113,385
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: 60/114,050
;  PRIOR FILING DATE: 1998-12-28
;  PRIOR APPLICATION NUMBER: 60/114,056
;  PRIOR FILING DATE: 1998-12-28
;  PRIOR APPLICATION NUMBER: 60/114,057
;  PRIOR FILING DATE: 1998-12-28
;  PRIOR APPLICATION NUMBER: 60/114,058
;  PRIOR FILING DATE: 1998-12-28
;  PRIOR APPLICATION NUMBER: 60/114,059
;  PRIOR FILING DATE: 1998-12-28
;  PRIOR APPLICATION NUMBER: 60/114,061
;  PRIOR FILING DATE: 1998-12-28
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
;   LENGTH: 963
;   TYPE: PRT
;   ORGANISM: Chlamydia pneumoniae
US-09-886-468-22

  Query Match             19.5%;  Score 989.5;  DB 10;  Length 963;
  Best Local Similarity   30.0%;  Pred. No. 3.9e-68;
  Matches  304;  Conservative  146;  Mismatches  416;  Indels  149;  Gaps   34;

Qy       9 LIGNSLSGLA-----REVPSRIFLMP----NSVPDPTKESLSNKISLTGDTHNLT----- 54
           | || :  ||      |:  |      |    |  |:|   |       | |  |::|
Db      30 LSGNEVPNLASCQMSRKDISAFHTSPSFRLNVTPEPLVSSFRPSNLLNGFGHDITQDITI 89

Qy      55 -----NCYLD-NLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAI 108
                |  :| |  |           :|  :    :    |     |  : | :|  :  ||||:
Db      90 TGNSINSVIDYNYHY---------EDGGILACKNL--FISENKGNLSFERNSSHSSGGAL 138

Qy     109 GYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNPN----------AAVNKIREGGAIHA 158
                     :|:             |  : | | |           |       ||||||
Db     139 ---------YSVRE------------C--WISKNQNYSFISNAASLATTTTSGFGGAIHA 175

Qy     159 QNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAI 218
           : || :|     |: | |  |||||    :  :::|    :   |  :: ::   || |
Db     176 LDSYITNNLGEGQFLDNVSKNRGGAIYVGVSLSITDNLGPIVIKKNQTLEDSSF--GGGI 233

Qy     219 YAGTSNSFESNNCDLFFINNACCAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEA 278
           :    |  | |  ::   :||   |   | |:  :  :   |:  :|    :: ||| :
Db     234 FCRAVN-IERNYQNIQINDNASGQGVVYFLPLGVIISSNKEIIEISNHSASSINTASGKL 292

Qy     279 SDGGAIKVTTRLDVTGN-RGRIFFSDNITKNYGGAIY-----APVVTLVDNGPTYFINNI 332
             ||   : | |    | :| ||  :| |    | :|     :  :|:     | ||||
Db     293 YPGGGGIMCTSLSHENNPKGLIF--NNKTAALSGGVYTRDLSSSKITV----RTAFINNS 346

Qy     333 ANNKGGAIYIDG----TSNSKISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVA 388
           | : |  | : |      |  :|||   |:|| | :        ||:|  |  |||:  |
Db     347 ATSGGALINLSGIGSTPQNFFLSADYGDILFNNNTI--------TSSSPQPGYRNAL-YA 397

Qy     389 SSSGEILLGAGSSQNLIFYDPIEVSNAGVS-VSFNKEADQTGSVVFSGATVNSADFHQRN 447
```

```
           :    : |||     ::||||||:        : || |    |:|:|||  |:|   :  |
Db     398 APGINLKLGARQGYKILFYDPIDHDQTTTDPIVFNYEPHHLGTVLFSGINVDSNATNPLN 457

Qy     448 LQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNSASNASI 507
             :|      |  | | ||| | ::    :||||:: ||| |::     | |     :||
Db     458 FLSKFSNSSRLERGVLAIEDRAAISCKTLSQTGGILRLGNAALIR--TKGPG-----SSI 510

Qy     508 TLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYE 567
             | :|| |||:| |  |  |: ||   : |: ||::| :||   |: ::|   :||:
Db     511 NFNAIAINLPSILQSEASAPKFWIYPTLTGSTYSEDTSSTITLSG-PLTFLNDENENPYD 569

Qy     568 STDLTHALSSQPMLSISEASDNQLRSDDMDFSGLNV------PHYGWQGLWTWGWAKTQD 621
           | ||     |:|   |       :     :| | | |       |||:||:|:  | :|
Db     570 SLDL-----SEPRKDIPPPLPPRCDCKKIDTSNLIVEAMNLDEHYGYQGIWSPYWMET-- 622

Qy     622 PEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLW--------GNMLL 673
             :: : | |:: |  || | : | | || |:|:     ||||||        |  :|
Db     623 ---TTTTSSTVPEQTNTNHRQLYVDWTPVGYRPNPERHGEFIANTLWQSAYNALLGIRIL 679

Qy     674 ATESLKN---SAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIA--G 728
             ::||     | |              |||::: |  ||   ||   ::||:|   |
Db     680 PPQNLKEHDLEASLQ------------GLGLLINQHNREGRKGFRNHTTGYAATTSAKTA 727

Qy     729 QTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGLYSYGDHN- 787
             |:||| |:| ::|  || : :  || ||         |   |: | |   ||||||:
Db     728 ARHSFSLGFAQMFSKTRERQSPSTTSSHNYFAGLRFDSLLFRDFISTGLSLGYSYGDHHM 787

Qy     788 -CHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHIL-TAPFLGALGIYSSLSHF 845
             ||  ||:    :|:  | : |:  ::  |    |    |   |   ||: |: :  | : |
Db     788 LCH--YTEILKGSSKAFFNNHTLVASL--DCTFLPARITRTLELQPFISAIALRCSQASF 843

Qy     846 TEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQL 905
           | | : | |  | || ::  ||| :  :  :   |  || |::| | |||: | : | |
Db     844 QETGDHIRKFHPKHPLTDLSSPIGFRSEWKTSHHIPMLWTTEISYVPTLYRKNPEMFTTL 903

Qy     906 LASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGE 960
           | | | |    :| | :::: ||   :|  | :||   |    ||||   ||  |
Db     904 LISNGTWTTQATPVSYNSVAAKIKNTSQLFSRVTLSLDYSAQVSSSTVGQYLKAE 958
```

```
RESULT 4
US-09-841-132-176
; Sequence 176, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
;  APPLICANT: Bhatia, Ajay
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C8
;  CURRENT APPLICATION NUMBER: US/09/841,132
;  CURRENT FILING DATE:  2001-04-23
;  NUMBER OF SEQ ID NOS: 599
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 176
;   LENGTH: 982
;   TYPE: PRT
;   ORGANISM: Chlamydia
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (1)...(982)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-09-841-132-176

  Query Match             12.7%;  Score 647;  DB 10;  Length 982;
  Best Local Similarity   25.4%;  Pred. No. 1.1e-41;
  Matches  257;  Conservative  156;  Mismatches  388;  Indels  212;  Gaps   43;

Qy      39 SLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAK 98
           :|    :: |  |:||    :|:|         |   |||::        ::  :|::
Db      54 NLLGSFTVLGRGHSLT---FENIR--------TSTNGAALS--------NSAADGLF--- 91
```

```
Qy      99 NLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNP-NAAV---------- 147
Db      92 --TIEGFKELSFSNCNSLLAVL-----PAATTNKGSQTPTTTSTPSNGTIYSKTDLLLLN 144

Qy     148 --------NKIR-EGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSC 198
Db     145 NEKFSFYSNLVSGDGGAIDAKSLTVQGISKLCVFQENTAQADGGACQVVTSFSAMANEAP 204

Qy     199 FLFMDNICIQTNTAG-KGGAIYA------GTSNSFE--------SNNCDLFFINNACCAG 243
Db     205 IAFVANV------AGVRGGGIAAVQDGQQGVSSSTSTEDPVVSFSRNTAVEFDGNVARVG 258

Qy     244 GAIFSPICSLTGNRGNIVFYNN--RCFKN----------------VETASSEASDGGAI 284
Db     259 GGIYS--------YGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQASNTSNNYGDGGAI 310

Qy     285 KVTTRLDVTGN---------RGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANN 335
Db     311 FCKNGAQAGSNNSGSVSFDGEGVVFFSSNVAAGKGGAIYAKKLSVANCGPVQFLRNIAND 370

Qy     336 KGGAIYIDGTSNSKISADRHAIIFNENI----VTNVTNANGTSTSANPPRRNAITVASSS 391
Db     371 -GGAIYLGESGELSLSADYGDIIFDGNLKRTAKENAADVNGVTVSS-----QAISMGSGG 424

Qy     392 GEILLGAGSSQNLIFYDPIEVSN-------AGVSVSFNKEADQTGSVVFSGATVNSADFH 444
Db     425 KITTLRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIVFAN---GSSTLY 481

Qy     445 QRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNS--A 502
Db     482 QN---------VTIEQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTLDFVTPQPPQQPPA 532

Qy     503 SNASITLKHIGLNLSSILKSGAEIPLLWVEPTN--NSNNYTADTAATFSLSDVKLS---L 557
Db     533 ANQLITLSNLHLSLSSLLANNA----VTNPPTNPPAQDSHPAVIGST-TAGSVTISGPIF 587

Qy     558 IDDYGNSPYESTD----------LTHALSSQPMLSISEASDNQLRSDDMDFSGLNVPHYG 607
Db     588 FEDLDDTAYDRYDWLGSNQKINVLKLQLGTKP--PANAPSDLTL--------GNEMPKYG 637

Qy     608 WQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTL 667
Db     638 YQGSWKLAW---------------DPNTANNGPYTLKATWTKTGYNPGPERVASLVPNSL 682

Qy     668 WGNMLLATESLKNSAELTPSDHPFW-GITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMI 726
Db     683 WGS-ILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYHDRDALGQGYRYISGGYSLG-- 739

Qy     727 AGQTHTFSLKFSQTYTKLNERYAKNNV--SSKNYSCQGEMLFSLQEGFLLTKLVGLYSYG 784
Db     740 -ANSYFGSSMFGLAFTEVFGR-SKDYVVCRSNHHACIGSVYLSTQQA-----LCGSYLFG 792

Qy     785 DHNCHHFYTQG-ENLTSQGTFRSQT--------MGGAVFFDLPMKPFGSTHILT--APFL 833
Db     793 DAFIRASYGFGNQHMKTSYTFAEESDVRWDNNCLAGEIGAGLPIVITPSKLYLNELRPFV 852

Qy     834 GALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPV 893
Db     853 QAEFSYADHESFTEEGDQARAFKS-GHLLNLSVPVGVKFDRCSST-HPNKYSFMAAYICD 910

Qy     894 LYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHG 946
Db     911 AYRTISGTETTLLSHQETWTTDAFHLARHGVVVRGSMYAS----LTSNIEVYG 959
```

```
RESULT 5
US-09-841-132-190
; Sequence 190, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
;  APPLICANT: Bhatia, Ajay
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C8
;  CURRENT APPLICATION NUMBER: US/09/841,132
;  CURRENT FILING DATE:  2001-04-23
;  NUMBER OF SEQ ID NOS: 599
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 190
;   LENGTH: 1006
;   TYPE: PRT
;   ORGANISM: Chlamydia
US-09-841-132-190

  Query Match             12.7%;  Score 647;  DB 10;  Length 1006;
  Best Local Similarity   25.4%;  Pred. No. 1.2e-41;
  Matches  257;  Conservative  156;  Mismatches  388;  Indels  212;  Gaps   43;

Qy      39 SLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAK 98
Db      78 NLLGSFTVLGRGHSLT---FENIR--------TSTNGAALS--------NSAADGLF--- 115

Qy      99 NLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNP-NAAV---------- 147
Db     116 --TIEGFKELSFSNCNSLLAVL-----PAATTNKGSQTPTTTSTPSNGTIYSKTDLLLLN 168

Qy     148 --------NKIR-EGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSC 198
Db     169 NEKFSFYSNLVSGDGGAIDAKSLTVQGISKLCVFQENTAQADGGACQVVTSFSAMANEAP 228

Qy     199 FLFMDNICIQTNTAG-KGGAIYA------GTSNSFE--------SNNCDLFFINNACCAG 243
Db     229 IAFVANV------AGVRGGGIAAVQDGQQGVSSSTSTEDPVVSFSRNTAVEFDGNVARVG 282

Qy     244 GAIFSPICSLTGNRGNIVFYNN--RCFKN----------------VETASSEASDGGAI 284
Db     283 GGIYS--------YGNVAFLNNGKTLFLNNVASPVYIAAKQPTSGQASNTSNNYGDGGAI 334

Qy     285 KVTTRLDVTGN---------RGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANN 335
Db     335 FCKNGAQAGSNNSGSVSFDGEGVVFFSSNVAAGKGGAIYAKKLSVANCGPVQFLRNIAND 394

Qy     336 KGGAIYIDGTSNSKISADRHAIIFNENI----VTNVTNANGTSTSANPPRRNAITVASSS 391
Db     395 -GGAIYLGESGELSLSADYGDIIFDGNLKRTAKENAADVNGVTVSS-----QAISMGSGG 448

Qy     392 GEILLGAGSSQNLIFYDPIEVSN-------AGVSVSFNKEADQTGSVVFSGATVNSADFH 444
Db     449 KITTLRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTGDIVFAN---GSSTLY 505

Qy     445 QRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNS--A 502
Db     506 QN---------VTIEQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTLDFVTPQPPQQPPA 556

Qy     503 SNASITLKHIGLNLSSILKSGAEIPLLWVEPTN--NSNNYTADTAATFSLSDVKLS---L 557
Db     557 ANQLITLSNLHLSLSSLLANNA----VTNPPTNPPAQDSHPAVIGST-TAGSVTISGPIF 611

Qy     558 IDDYGNSPYESTD----------LTHALSSQPMLSISEASDNQLRSDDMDFSGLNVPHYG 607
Db     612 FEDLDDTAYDRYDWLGSNQKINVLKLQLGTKP--PANAPSDLTL--------GNEMPKYG 661

Qy     608 WQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTL 667
Db     662 YQGSWKLAW---------------DPNTANNGPYTLKATWTKTGYNPGPERVASLVPNSL 706

Qy     668 WGNMLLATESLKNSAELTPSDHPFW-GITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMI 726
Db     707 WGS-ILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYHDRDALGQGYRYISGGYSLG-- 763

Qy     727 AGQTHTFSLKFSQTYTKLNERYAKNNV--SSKNYSCQGEMLFSLQEGFLLTKLVGLYSYG 784
Db     764 -ANSYFGSSMFGLAFTEVFGR-SKDYVVCRSNHHACIGSVYLSTQQA-----LCGSYLFG 816
```

```
Qy    785 DHNCHHFYTQG-ENLTSQGTFRSQT--------MGGAVFFDLPMKPFGSTHILT--APFL 833
Db    817 DAFIRASYGFGNQHMKTSYTFAEESDVRWDNNCLAGEIGAGLPIVITPSKLYLNELRPFV 876

Qy    834 GALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPV 893
Db    877 QAEFSYADHESFTEEGDQARAFKS-GHLLNLSVPVGVKFDRCSST-HPNKYSFMAAYICD 934

Qy    894 LYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHG 946
Db    935 AYRTISGTETTLLSHQETWTTDAFHLARHGVVVRGSMYAS----LTSNIEVYG 983
```

RESULT 6
US-09-452-380-4
; Sequence 4, Application US/09452380
; Patent No. US20020094340A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, Andrew D.
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: WANG, Joe
; APPLICANT: DUNN, Pamela
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 032931/0216
; CURRENT APPLICATION NUMBER: US/09/452,380
; CURRENT FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: US 60/132,272
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: US 60/113,439
; PRIOR FILING DATE: 1998-12-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 925
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-452-380-4

Query Match 12.7%; Score 645; DB 10; Length 925;
Best Local Similarity 27.0%; Pred. No. 1.5e-41;
Matches 262; Conservative 146; Mismatches 392; Indels 172; Gaps 44;

```
Qy     75 GAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGY------------ASPNSPTVEIR 121
Db     43 GTTYSLLSDVSFQNAGALGIPLASGCFLEAGGDLTFQGNQHALKFAFINAGSSAGTVAST 102

Qy    122 DTIGPVIFENN-------TCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMK 174
Db    103 SAADKNLLFNDFSRLSIISCPSLLLSPTGQCALKSV--------GNLSLTGNSQII-FTQ 153

Qy    175 NFSYVRGGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYA-GTSNSFESNNCDL 233
Db    154 NFSSDNGGVINTKN-FLLSGTSQFASFSRN---QAFTGKQGGVVYATGTITIENSPGIVS 209

Qy    234 FFINNACCAGGAIFS-PICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTT---R 289
Db    210 FSQNLAKGSGGALYSTDNCSITDN-FQVIFDGNSAWEAAQ------AQGGAICCTTTDKT 262

Qy    290 LDVTGNRGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNK-----GGAIYIDG 344
Db    263 VTLTGNK-NLSFTNNTALTYGGAISGLKVSISAGGPTLFQSNISGSSAGQGGGGAINIAS 321

Qy    345 TSNSKISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNL 404
Db    322 AGELALSATSGDITFNNNQVTN----GSTST------RNAINIIDTAKVTSIRAATGQSI 371

Qy    405 IFYDPIEVSNAGVSVS----------FNKEADQTGSVVFSGATVNSAD-FHQRNLQTKTP 453
Db    372 YFYDPI--TNPGTAASTDTLNLNLADANSEIEYGGAIVFSGEKLSPTEKAIAANVTSTIR 429

Qy    454 APLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNSASNASITLKHIG 513
Db    430 QPAVLARGDLVLRDGVTVTFKDLTQSPGSRILMDG--------GTTLSAKEANLSLNGLA 481

Qy    514 LNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTH 573
Db    482 VNLSSL-----------DGTNKAALKTEAADKNISLSGT-IALIDTEG-SFYENHNLKS 527

Qy    574 ALSSQPMLSISEASDN-QLRSDDMDFSGLNVP--HYGWQGLWTWGWAKTQDPEPASSATI 630
Db    528 A-STYPLLELTTAGANGTITLGALSTLTLQEPETHYGYQGNWQLSWAN------ATSSKI 580

Qy    631 TDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSDHP 690
Db    581 GS-----------INWTRTGYIPSPERKSNLPLNSLWGN-FIDIRSINQLIETKSSGEP 627

Qy    691 F----WGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFSLKFSQTYTKLNE 746
Db    628 FERELW---LSGIANFFYRDSMPTRHGFRHISGGYALGITATTPAEDQLTFA--FCQLFA 682

Qy    747 RYAKNNVSSKN---------YSCQGEMLFSLQEGFL-------------LTKLVGL---- 780
Db    683 R-DRNHITGKNHGDTYGASLYFHHTEGLFDI-ANFLWGKATRAPWVLSEISQIIPLSFDA 740

Qy    781 ---YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHIL--TAPFLGA 835
Db    741 KFSYLHTDNHMKTYYT--DNSIIKGSWRNDAFCADLGASLPF-VISVPYLLKEVEPFVKV 797

Qy    836 LGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVL- 894
Db    798 QYIYAHQQDFYERHAEGRAFN-KSELINVEIPIGVT---FERDSKSEKGTYDLTLMYILD 853

Qy    895 -YRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTF 953
Db    854 AYRRNPKCQTSLIASDANWMAYGTNLARQGFSVRAANHFQVNPHMEIFGQF-AFEVRSSS 912

Qy    954 CNYLNGEIALRF 965
Db    913 RNY-NTNLGSKF 923

RESULT 7
US-09-452-380-3
; Sequence 3, Application US/09452380
; Patent No. US20020094340A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, Andrew D.
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: WANG, Joe
; APPLICANT: DUNN, Pamela
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 032931/0216
; CURRENT APPLICATION NUMBER: US/09/452,380
; CURRENT FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: US 60/132,272
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: US 60/113,439
; PRIOR FILING DATE: 1998-12-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 936
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-452-380-3

Query Match 12.7%; Score 645; DB 10; Length 936;
Best Local Similarity 27.0%; Pred. No. 1.5e-41;
Matches 262; Conservative 146; Mismatches 392; Indels 172; Gaps 44;

```
Qy   75 GAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGY------------ASPNSPTVEIR 121
Db   54 GTTYSLLSDVSFQNAGALGIPLASGCFLEAGGDLTFQGNQHALKFAFINAGSSAGTVAST 113
```
```

```
Qy    122 DTIGPVIFENN-------TCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMK 174
Db    114 SAADKNLLFNDFSRLSIISCPSLLLSPTGQCALKSV--------GNLSLTGNSQII-FTQ 164

Qy    175 NFSYVRGGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYA-GTSNSFESNNCDL 233
Db    165 NFSSDNGGVINTKN-FLLSGTSQFASFSRN---QAFTGKQGGVVYATGTITIENSPGIVS 220

Qy    234 FFINNACCAGGAIFS-PICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTT---R 289
Db    221 FSQNLAKGSGGALYSTDNCSITDN-FQVIFDGNSAWEAAQ------AQGGAICCTTTDKT 273

Qy    290 LDVTGNRGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNK-----GGAIYIDG 344
Db    274 VTLTGNK-NLSFTNNTALTYGGAISGLKVSISAGGPTLFQSNISGSSAGQGGGGAINIAS 332

Qy    345 TSNSKISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNL 404
Db    333 AGELALSATSGDITFNNNQVTN----GSTST------RNAINIIDTAKVTSIRAATGQSI 382

Qy    405 IFYDPIEVSNAGVSVS----------FNKEADQTGSVVFSGATVNSAD-FHQRNLQTKTP 453
Db    383 YFYDPI--TNPGTAASTDTLNLNLADANSEIEYGGAIVFSGEKLSPTEKAIAANVTSTIR 440

Qy    454 APLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNSASNASITLKHIG 513
Db    441 QPAVLARGDLVLRDGVTVTFKDLTQSPGSRILMDG--------GTTLSAKEANLSLNGLA 492

Qy    514 LNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTH 573
Db    493 VNLSSL------------DGTNKAALKTEAADKNISLSGT-IALIDTEG-SFYENHNLKS 538

Qy    574 ALSSQPMLSISEASDN-QLRSDDMDFSGLNVP--HYGWQGLWTWGWAKTQDPEPASSATI 630
Db    539 A-STYPLLELTTAGANGTITLGALSTLTLQEPETHYGYQGNWQLSWAN------ATSSKI 591

Qy    631 TDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSDHP 690
Db    592 GS------------INWTRTGYIPSPERKSNLPLNSLWGN-FIDIRSINQLIETKSSGEP 638

Qy    691 F----WGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFSLKFSQTYTKLNE 746
Db    639 FERELW---LSGIANFFYRDSMPTRHGFRHISGGYALGITATTPAEDQLTFA--FCQLFA 693

Qy    747 RYAKNNVSSKN---------YSCQGEMLFSLQEGFL-------------LTKLVGL---- 780
Db    694 R-DRNHITGKNHGDTYGASLYFHHTEGLFDI-ANFLWGKATRAPWVLSEISQIIPLSFDA 751

Qy    781 ---YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHIL--TAPFLGA 835
Db    752 KFSYLHTDNHMKTYYT--DNSIIKGSWRNDAFCADLGASLPF-VISVPYLLKEVEPFVKV 808

Qy    836 LGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVL- 894
Db    809 QYIYAHQQDFYERHAEGRAFN-KSELINVEIPIGVT---FERDSKSEKGTYDLTLMYILD 864

Qy    895 -YRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTF 953
Db    865 AYRRNPKCQTSLIASDANWMAYGTNLARQGFSVRAANHFQVNPHMEIFGQF-AFEVRSSS 923

Qy    954 CNYLNGEIALRF 965
Db    924 RNY-NTNLGSKF 934


RESULT 8
US-10-023-437-57
; Sequence 57, Application US/10023437
; Publication No. US20020183272A1
; GENERAL INFORMATION:
;  APPLICANT: JOHNSTON, STEPHEN A.
;  APPLICANT:  STEMKE-HALE, KATHERINE
;  APPLICANT:  SYKES, KATHRYN F.
;  APPLICANT:  KALTENBOECK, BERNHARD
;  TITLE OF INVENTION: METHODS AND compositions for Vaccination COMPRISING NUCLEIC AC
;  TITLE OF INVENTION:  AND/OR POLYPEPTIDE SEQUENCES OF CHLAMYDIA
;  FILE REFERENCE: UTSD:736US
;  CURRENT APPLICATION NUMBER: US/10/023,437
;  CURRENT FILING DATE:  2001-12-17
;  PRIOR APPLICATION NUMBER: 60/225,839
;  PRIOR FILING DATE: 2000-12-15
;  NUMBER OF SEQ ID NOS: 69
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
;   LENGTH: 926
;   TYPE: PRT
;   ORGANISM: Chlamydia psittaci
US-10-023-437-57

  Query Match            12.5%;  Score 636;  DB 9;  Length 926;
  Best Local Similarity   24.0%;  Pred. No. 7.3e-41;
  Matches  241;  Conservative  157;  Mismatches  365;  Indels  240;  Gaps   41;

Qy     41 SNKISLTGDTHNL--TNCYLDNLRYILAILQKTPNEGA-----------AVTITDY--L 84
Db     82 ADNLTFKGNNHSLSITNA---------------NAGANPAGINVNTADKILTLTDFSKL 125

Qy     85 SFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNPN 144
Db    126 SFKECPSSLVNTGKG-AMKSGGALNLANNAS----------ILFDQN-----YSAEN-- 166

Qy    145 AAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSCFLFMDN 204
Db    167 --------GGAISCKAFSLTGSSKEISFTTNSTAKKGGAIAATGIAHLSDNQGTIRFSGN 218

Qy    205 ICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNACC-----AGGAIFSPICSLTGNRGN 259
Db    219 TAVNS-----GGAVYSEASMTIAGNN-HVAFSNNAVSGSSDGCGGAIH---CSKTGSAPT 269

Qy    260 IVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYAPVVT 319
Db    270 LTIRDNKV-----------------------------LIFEENTSSAKGGAIYTDKLI 298

Qy    320 LVDNGPTYFINN---IANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTSTS 376
Db    299 LTSGGPTAFINNKVTHATPKGGAIGIAANGECSLTAEHGDITFDNNLMATQDNAT----- 353

Qy    377 ANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEV-SNAGVSVSFNK-EADQT--GSVV 432
Db    354 ---IKRNAINIEGNGKFVNLRAASGKTISFYDPITVEGNAADLLTLNKAEGDKTYNGRII 410

Qy    433 FSGATVNS-----ADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGN 487
Db    411 FSGEKLTEEQAAVAD----NLKTTFTQPITLAAGELVLRSGVEVEAKTVVQTAGSLILMD 466

Qy    488 GAVLSCYKNGAGNSASNASITLKHIGLNLSSI----------LKSGAEIPLLW----VEP 533
Db    467 A--------GTKLSAKTEDATLTNLAINPNTLDGKKFAVVDAVAAGKNVTLSGAIGVIDP 518

Qy    534 TNN--SNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDL-THALSSQPMLSISEASDNQ 590
Db    519 TGKFYENHKLNDTLA---LGGIQLS-----GKGSVTTTNVPSHVV-------------- 555

Qy    591 LRSDDMDFSGLNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPA 650
Db    556 ---------GVAETHYGYQGNWSVSWVKDNNSDPKT-------------QTAIFTWNKT 592

Qy    651 GYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSD--HPFWGITGGGLGMMVYQDPR 708
Db    593 GYVPNPERRAPLVLNSLWGS-FIDLRSIQDVLERSVDSILETRRGLWVSGIGNFFHKDRN 651

Qy    709 ENHPGFHMRSSGYSAGMIAGQTH--TFSLKFSQTYTKLNERYAKNNVSS---------- 755
Db    652 AENRKFRHISSGYVLGATTNTSREDSLSVAFCQLFAKDKDYLVSKNAANVYAGSVYYQHV 711

Qy    756 ------------KNYSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQGENLTSQGT 803
```

```
Db    712 SKFDDLTRLFNGPNTCCSG---FSKEIPIFLDAQI-TYCHTANNMTTSYTDYPEV--KGS 765

Qy    804 FRSQTMGGAVFFDLPMKPFGSTHILT-APFLGALGIYSSLSHFTEVGAYPRSFSTKTPLI 862
Db    766 WGNDTLGLTLSTSVPIPVFSSSIFDSYAPFAKLQVVYAHQDDFKEPTTEGRVFES-SDLL 824

Qy    863 NVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRH 922
Db    825 NVSVPIGIKFEKLSYGER-SAYDLTLMYIPDVYRHNPSCMTGLAINDVSWLTTATNLARQ 883

Qy    923 AMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNY---LNGEIA 962
Db    884 AFIVRAGNHIALTSGVEMFSQF-GFELRSSSRNYNVDLGAKVA 925


RESULT 9
US-09-841-132-321
; Sequence 321, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
;  APPLICANT: Bhatia, Ajay
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C8
;  CURRENT APPLICATION NUMBER: US/09/841,132
;  CURRENT FILING DATE:  2001-04-23
;  NUMBER OF SEQ ID NOS: 599
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 321
;   LENGTH: 715
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis
US-09-841-132-321

  Query Match            11.9%;  Score 605.5;  DB 10;  Length 715;
  Best Local Similarity   27.2%;  Pred. No. 1.1e-38;
  Matches  208;  Conservative  120;  Mismatches  277;  Indels  159;  Gaps   30;

Qy    302 SDNITKNYGGAIYA-PV-----------VTLVDNGPTYFIN-NIANNKGGAIYIDGTSNS 348
Db     11 SDNFQLSQGGQGFAIPIGQAMAIAGQIKLPTVHIGPTAFLGLGVVDNNGNGARVQRVVGS 70

Qy    349 KISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYD 408
Db     71 APAASLG--ISTGDVITAVDGA------------PINSATAMADALNG---------HH 106

Qy    409 PIEVSNAGVSVSFNKEA--DQTGSVVFSGATVNSADF-----HQRNLQTKT--------- 452
Db    107 PGDV----ISVTWQTKSGGTRTGNVTL--AEGPPAEFCRYPSHWRPLDIRTLMGKEHNYI 160

Qy    453 -PAPLTLSNGFLCIEDHAQLTVNR--FTQT-GGVVSLGNGAVLSCYKNGAGNSASNASIT 508
Db    161 KEAPTTLKFGTLAIEDDAELEIFNIPFTQNPTSLLALGSGATLTVGKHG--------KLN 212

Qy    509 LKHIGLNLSSILKSGAEIPLLWVEPTN-NSNNYTADT-AATFSLS------DVKLSLIDD 560
Db    213 ITNLGVILPIILKEGKSPPCIRVNPQDMTQNTGTGQTPSSTSSISTPMIIFNGRLSIVDE 272

Qy    561 YGNSPYESTDLTHALSSQPMLSISEASDNQLRSD---DMDFSGLNVPHYGWQGLWTWGWA 617
Db    273 NYESVYDSMDLSRGKAEQLILSIETTNDGQLDSNWQSSLNTSLLSPPHYGYQGLWTPNWI 332

Qy    618 KT--------QDPEPASSATITDPQKANR------------------------------ 638
Db    333 TTTYTITLNNNSSAPTSATSIAEQKKTSETFTPSNTTTASIPNIKASAGSGSGSASNSGE 392

Qy    639 ---FHRTLLLTWLPAGYVPSPKHRSPLIANTL---WGNMLLATESLKNSAELTPSDHPFW 692
Db    393 VTITKHTLVVNWAPVGYIVDPIRRGDLIANSLVHSGRNMTMGLRSLL-------PDNSWF 445

Qy    693 GITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIA--GQTHTFSLKFSQTYTKLNERYAK 750
Db    446 ALQGAATTLFTKQQKRLSYHGYSSASKGYTVSSQASGAHGHKFLLSFSQSSDKMKEKETN 505

Qy    751 NNVSSKNY---SCQGEMLFSLQEGFLLTKLVG--LYSYGDHNCHHFYTQGENLTSQGTFR 805
Db    506 NRLSSRYYLSALCFEHPMFD------RIALIGAAACNYGTHNMRSFY--GTKKSSKGKFH 557

Qy    806 SQTMGGAVFFDL----PMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPL 861
Db    558 STTLGASLRCELRDSMPLR-----SIMLTPFAQALFSRTEPASIRESGDLARLFTLEQAH 612

Qy    862 INVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSSR 921
Db    613 TAVVSPIGIKGAYSSDTWPTLSWEMELAYQPTLYWKRPLLNTLLIQNNGSWVTTNTPLAK 672

Qy    922 HAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
Db    673 HSF-YGRGSHSLKFSHLKLFANYQAEVATSTVSHYINAGGALVF 715


RESULT 10
US-10-007-693-95
; Sequence 95, Application US/10007693
; Patent No. US20020146776A1
; GENERAL INFORMATION:
;  APPLICANT: Bhatia, Ajay
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
;  TITLE OF INVENTION:  AND DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.515C2
;  CURRENT APPLICATION NUMBER: US/10/007,693
;  CURRENT FILING DATE:  2001-12-05
;  NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 95
;   LENGTH: 1016
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis serovar D
US-10-007-693-95

  Query Match            11.7%;  Score 595;  DB 12;  Length 1016;
  Best Local Similarity   25.4%;  Pred. No. 1.2e-37;
  Matches  261;  Conservative  149;  Mismatches  404;  Indels  214;  Gaps   49;

Qy     42 NKISLTGDTHNLT-----NCYLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYF 96
Db     69 DNLTITGQNHTLSFTDSQGPVLQNYAFISA--------GETLTLKDFSS--------LMF 112

Qy     97 AKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNT-------------------CCRP 137
Db    113 SKNVSCGEKGMI-----SGKTVSISGA-GEVIFWDNSVGYSPLSIVPASTPTPPAPAPAP 166

Qy    138 FTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQS 197
Db    167 AASSSLSPTVSDARKGSIFSVET-----SLEISGVKKGVMF--------------DNNAG 207

Qy    198 CF--LFMDNICIQTNTAGKGGAIYAGTSNSFESNNCD---LFFINNACCAGGAIFSPICS 252
Db    208 NFGTVFRGN---SNNNAGSGGS-GSATTPSFTVKNCKGKVSFTDNVASCGGGVVYKGTVL 263

Qy    253 LTGNRGNIVFYNNRCFKNVETASSEASD--------GGAI-KVTTRLDVTGNRGRIFFSD 303
Db    264 FKDNEGGIFFRGNTAYDDLGILAATSRDQNTETGGGGGVICSPDDSVKFEGNKGSIVFDY 323

Qy    304 NITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIY---IDGTSNS-------KISAD 353
Db    324 NFAKGRGGSILTKEFSLVADDSVVFSNNTAEKGGGAIYAPTIDISTNGGSILFERNRAAE 383

Qy    354 RHAIIFNE-------NIVTNVTNA----NGTSTSANPPRRNAITVASSSGEILLGAGSSQ 402
Db    384 GGAICVSEASSGSTGNLTLSASDGDIVFSGNMTSDRPGERSAARILSDGTTVSLNASGLS 443

Qy    403 NLIFYDPIEVSN--AGVS------------VSFNKEADQTGSVVFSGATVNSADFHQ-RN 447
```

```
Db    444 KLIFYDPVVQNNSAAGASTPSPSSSSMPGAVTINQSGN--GSVIFTAESLTPSEKLQVLN 501

Qy    448 LQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNSASNASI 507

Db    502 STSNFPGALTVSGGELVVTEGATLTTGTITATSGRVTLGSGASLSAVAGAANN---NYTC 558

Qy    508 TLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYE 567

Db    559 TVSKLGIDLESFLTPNYKTAILGADGT-----VTVNSGSTLDL--VMESEAEVYDN-PLF 610

Qy    568 STDLTHALSSQPMLSISEAS-DNQLRSDDMDFSGLNVPHYGWQGLWTWGWAKTQDPEPAS 626

Db    611 VGSLT-----IPFVTLSSSSASNGVTKNSVTINDADAAHYGYQGSWSADWTK---PPLAP 662

Qy    627 SATITDPQKANRFHRTLLLTWLPAG----YVPSPKHRSPLIANTLW--GNMLLA-TESLK 679

Db    663 DAKGMVPPNTN---NTLYLTWRPASNYGEYRLDPQRKGELVPNSLWVAGSALRTFTNGLK 719

Qy    680 -----NSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTH--- 731

Db    720 EHYVSRDVGFVASLH--------ALGDYILNYTQDDRDGFLARYGGFQA---TAASHYEN 768

Qy    732 --TFSLKFSQTY--TKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGLY--SYGD 785

Db    769 GSIFGVAFGQLYGQTKSRMYYSKDAGNMTMLSCFGRSYVDIKG----TETVMYWETAYG- 823

Qy    786 HNCHHFYTQGENLTSQGTFRSQ---------TMGGA----VFFDLPMKPFGSTHILTA-- 830

Db    824 YSVHRMHTQYFNDKTQKFDHSKCHWHNNNYYAFVGAEHNFLEYCIPTRQFARDYELTGFM 883

Qy    831 PFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQR--PQAWTVEL 888

Db    884 RFEMAGGWSSST---RETGSLTRYFARGSG-HNMSLPIGIVAHAVSHVRRSPPSKLTLNM 939

Qy    889 AYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFY 948

Db    940 GYRPDIWRVTPHCNMEIIANGVKTPIQGSPLARHAFFLEVHD--------TLYIHHFG-- 989

Qy    949 SSSTFCNY 956

Db    990 --RAYMNY 995


RESULT 11
US-10-007-693-65
; Sequence 65, Application US/10007693
; Patent No. US20020146776A1
; GENERAL INFORMATION:
;  APPLICANT: Bhatia, Ajay
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
;  TITLE OF INVENTION:  AND DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.515C2
;  CURRENT APPLICATION NUMBER: US/10/007,693
;  CURRENT FILING DATE:  2001-12-05
;  NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 65
;   LENGTH: 978
;   TYPE: PRT
;   ORGANISM: Chlamydia pneumoniae
US-10-007-693-65

  Query Match             11.5%;  Score 583;  DB 12;  Length 978;
  Best Local Similarity   26.1%;  Pred. No. 9.9e-37;
  Matches  263;  Conservative  137;  Mismatches  411;  Indels  196;  Gaps   45;

Qy     41 SNKISLTGDTHNL--TNCYL-DNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFA 97

Db     67 SSNLSLLGKGLSLTFTSCQAPTNSNYALL--------SAAETLT-FKNF-----SSINFT 112

Qy     98 KNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENN----TCCRPFTSSNPNAAVNKIREG 153

Db    113 GNQSTGLGGLI-YGKD-----IVFQSIKDLIFTTNRVAYSPASVTTSATP--AITTVTTG 164

Qy    154 G-------AIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVS--ENQSCFLFMDN 204

Db    165 ASALQPTDSLTVEN--ISQSIKFFGNLANF----GSAISSSPTAVVKFINNTATMSFSHN 218

Qy    205 ICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNACC----------------AGGAIFS 248

Db    219 F-----TSSGGGVIYGGSSLLFENNSGCIIFTANSCVNSLKGVTPSSGTYALGSGGAICI 273

Qy    249 P--ICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNIT 306

Db    274 PTGTFELKNNQGKCTFSYN----------GTPNDAGAIYAET-CNIVGNQGALLLDSNTA 322

Qy    307 KNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYI--------DGTSNSKISADRHAII 358

Db    323 ARNGGAICAKVLNIQGRGPIEFSRNRA-EKGGAIFIGPSVGDPAKQTSTLTILASEGDIA 381

Qy    359 FNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPI-------E 411

Db    382 FQGNML-----------NTKPGIRNAITVEAGGEIVSLSAQGGSRLVFYDPITHSLPTTS 430

Qy    412 VSNAGVSVSFNKEADQTGSVVFSGATVNSAD-FHQRNLQTKTPAPLTLSNGFLCIEDHAQ 470

Db    431 PSNKDITINAN---GASGSVVFTSKGLSSTELLLPANTTTILLGTVKIASGELKITDNAV 487

Qy    471 LTVNRF-TQTGGVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGAEIPLL 529

Db    488 VNVLGFATQGSGQLTLGSGGTLGLATPTGAPAA--VDFTIGKLAFDPFSFLKRD------ 539

Qy    530 WVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDN 589

Db    540 FVSASVNAG--TKNVTLTGAL------VLDE-----HDVTDLYDMVSLQSPVAIPIAV-- 584

Qy    590 QLRSDDMDFSGLNV-------------PHYGWQGLWTWGWAK-TQDPEPASSATITDPQK 635

Db    585 --------FKGATVTKTGFPDGEIATPSHYGYQGKWSYTWSRPLLIPAPDGGFPGGPSPS 636

Qy    636 ANRFHRTLLLTW-----LPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSA------EL 684

Db    637 AN----TLYAVWNSDTLVRSTYILDPERYGEIVSNSLW------ISFLGNQAFSDILQDV 686

Qy    685 TPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTH--TFSLKFSQTYT 742

Db    687 LLIDHPGLSITAKALGAYVEHTPRQGHEGFSGRYGGYQAALSMNYTDHTTLGLSFGQLYG 746

Qy    743 KLNERYAKNNVSSKNY--SCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQGENL-T 799

Db    747 KTNANPYDSRCSEQMYLLSFFGQFPIVTQKSEALISWKAAYGYSKNHLNTTYLRPDKAPK 806

Qy    800 SQGTFRSQTMGGAVFFDLPMKPFGSTHILTAP---------FLGALGIYSSLSHFTEVGA 850

Db    807 SQGQWHNNSYYVLISAE---HPFLNWCLLTRPLAQAWDLSGFISAEFLGGWQSKFTETGD 863

Qy    851 YPRSFSTKTPLINVLVPIGVKGSFMNATQR-PQAWTVELAYQPVLYRQEPGIATQLLASK 909

Db    864 LQRSFS-RGKGYNVSLPIGCSSQWFTPFKKAPSTLTIKLAYKPDIYRVNPHNIVTVVSNQ 922

Qy    910 GIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNY 956

Db    923 ESTSISGANLRRHGLFVQI-HDVVDLTEDTQAFLNYTFDGKNGFTNH 968


RESULT 12
US-09-886-468-19
; Sequence 19, Application US/09886468
; Patent No. US20020037293A1
; GENERAL INFORMATION:
;  APPLICANT: Aventis Pasteur Limited
;  TITLE OF INVENTION: Chlamydia antigens and corresponding DNA fragments and uses th
;  FILE REFERENCE: 77813-5
;  CURRENT APPLICATION NUMBER: US/09/886,468
;  CURRENT FILING DATE:  1999-12-23
```

```
;  PRIOR APPLICATION NUMBER: 60/113,280
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: 60/113,281
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: 60/113,282
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: 60/113,283
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: 60/113,284
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: 60/113,285
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: 60/113,385
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: 60/114,050
;  PRIOR FILING DATE: 1998-12-28
;  PRIOR APPLICATION NUMBER: 60/114,056
;  PRIOR FILING DATE: 1998-12-28
;  PRIOR APPLICATION NUMBER: 60/114,057
;  PRIOR FILING DATE: 1998-12-28
;  PRIOR APPLICATION NUMBER: 60/114,058
;  PRIOR FILING DATE: 1998-12-28
;  PRIOR APPLICATION NUMBER: 60/114,059
;  PRIOR FILING DATE: 1998-12-28
;  PRIOR APPLICATION NUMBER: 60/114,061
;  PRIOR FILING DATE: 1998-12-28
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
;   LENGTH: 922
;   TYPE: PRT
;   ORGANISM: Chlamydia pneumoniae
US-09-886-468-19

  Query Match             11.2%;  Score 569.5;  DB 10;  Length 922;
  Best Local Similarity   24.5%;  Pred. No. 1e-35;
  Matches  236;  Conservative  151;  Mismatches  354;  Indels  221;  Gaps   45;

Qy      45 SLTGDTHNLTNCYLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYF---AKNLT 101
           |||||  :::|  :||     : | |     |   :|     |     |:||   :   |
Db      57 SLTGDV-SISN--VDN-----SALNK-----ACFNVTSGSVTFAGNHHGLYFNNISSGTT 103

Qy     102 PESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRP---------FTSSNPNAAVNKIRE 152
           |  ||:                          ||:          |:: :   :   |:|
Db     104 KE--GAV----------------------LCCQDPQATARFSGFSTLSFIQSPGDIKE 137

Qy     153 GGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSCFLFMDNICIQTNTA 212
           | ::::|  :  |: || | :| |  :|||||  ||  :|          |::    | |
Db     138 QGCLYSKNALMLLNNYVVRFEQNQSKTKGGAISGANVTIVGN-------YDSVSFYQNAA 190

Qy     213 GKGGAIYAGTSNSFESNNCDLFFINNAC--CAGGAIFSPICSLTGNRGNIVFYNNR--CF 268
             ||||::        |  :: | |     :|||::|          |:|    |    |
Db     191 TFGGAIHSSGPLQIAVNQAEIRFAQNTAKNGSGGALYSD--------GDIDIDQNAYVLF 242

Qy     269 KNVETASSEASDGGAIKVTTRLDVTGNRGR---IFFSDN----ITKNY----GGAIYAPV 317
           :  |  ::    |||:     |  :|:      : ||||      :|:    ||||||
Db     243 RENEALTTAIGKGGAV---CCLPTSGSSTPVPIVTFSDNKQLVFERNHSIMGGGAIYARK 299

Qy     318 VTLVDNGPTYFINNI----ANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGT 373
           :::   ||| |||||    : | |||| ||      :||::  | |             |
Db     300 LSISSGGPTLFINNISYANSQNLGGAIAIDTGGEISLSAEKGTITF----------QGN 348

Qy     374 STSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVSFN-----KEADQT 428
           ||  |  | | : ::  : | | :  :: |||||  | |  |   |      |  : |
Db     349 RTSL--PFLNGIHLLQNAKFLKLQARNGYSIEFYDPI-TSEADGSTQLNINGDPKNKEYT 405

Qy     429 GSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGG---VVSL 485
           |:::|||    : |   |: ::  |  : || |:| |:: |::||::|||: |   |: |
Db     406 GTILFSGEKSLAND--PRDFKSTIPQNVNLSAGYLVIKEGAEVTVSKFTQSPGSHLVLDL 463

Qy     486 GNGAVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTA 545
           |   :          ||   | :  : ::: |: |       |  | :|
Db     464 GTKLI----------ASKEDIAITGLAIDIDSLSSSSTAA----VIKANTANK------ 502

Qy     546 ATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSGLNVPH 605
             |::|  : ||   ||: ||   : ::  : |:||:   :   :     ||  :: ||
Db     503 -QISVTD-SIELISPTGNA-YEDLRMRNS-QTFPLLSLEPGAGGSVTVTAGDFLPVS-PH 557

Qy     606 YGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIAN 665
           ||:|| |   |  | :             |   |       |    | | |:    |: |
Db     558 YGFQGNWKLAWTGTGN------------KVGEFF------WDKINYKPRPEKEGNLVPN 598

Qy     666 TLWGN------MLLATESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSS 719
           ||||      ::   |:  :| :   :|   |     |:|   :    |::  :   |
Db     599 ILWGNAVDVRSLMQVQETHASSLQ---TDRGLW---IDGIGNFFHVSASEDNIRYRHNSG 652

Qy     720 GY--SAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKL 777
           ||  |      |  |: ||| ::: :: || :|   : |   |  |:    | | |
Db     653 GYVLSVNNEITPKHYTSMAFSQLFSR-DKDYAVSNNEYRMY--LGSYLYQ-----YTTSL 704

Qy     778 VGLYSYGDHN----------------------CHHFYTQGENLTSQGTFR--------- 805
             :: |   |                      | : :   :  |    |
Db     705 GNIFRYASRNPNVNVGILSRRFLQNPLMIFHFLCAYGHATNDMKTDYANFPMVKNSWRNN 764

Qy     806 --SQTMGGAVFFDLPMKPFGSTHIL--TAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPL 861
             :   ||:    :|:  | :  :     ||:    :|:    | |  |  | ||  : |
Db     765 CWAIECGGS----MPLLVFENGRLFQGAIPFMKLQLVYAYHGDFKETTADGRRFSNGS-L 819

Qy     862 INVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSSR 921
           :: ||:|::   : |  :  :    :| | ::|::|     |: |   |   :  ||
Db     820 TSISVPLGIRFEKL-ALSQDVLYDFSFSYIPDIFRKDPSCEAALVISGDSWLVPAAHVSR 878

Qy     922 HA 923
           ||
Db     879 HA 880


RESULT 13
US-10-023-437-23
; Sequence 23, Application US/10023437
; Publication No. US20020183272A1
; GENERAL INFORMATION:
;  APPLICANT: JOHNSTON, STEPHEN A.
;  APPLICANT:  STEMKE-HALE, KATHERINE
;  APPLICANT:  SYKES, KATHRYN F.
;  APPLICANT:  KALTENBOECK, BERNHARD
;  TITLE OF INVENTION: METHODS AND compositions for Vaccination COMPRISING NUCLEIC AC
;  TITLE OF INVENTION:  AND/OR POLYPEPTIDE SEQUENCES OF CHLAMYDIA
;  FILE REFERENCE: UTSD:736US
;  CURRENT APPLICATION NUMBER: US/10/023,437
;  CURRENT FILING DATE:  2001-12-17
;  PRIOR APPLICATION NUMBER: 60/225,839
;  PRIOR FILING DATE: 2000-12-15
;  NUMBER OF SEQ ID NOS: 69
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
;   LENGTH: 839
;   TYPE: PRT
;   ORGANISM: Chlamydia psittaci
US-10-023-437-23

  Query Match             10.4%;  Score 531;  DB 9;  Length 839;
  Best Local Similarity   25.6%;  Pred. No. 8.1e-33;
  Matches  232;  Conservative  129;  Mismatches  347;  Indels  198;  Gaps   44;

Qy     167 HDVVGFMKNFSYVRGGAISTAN------------------TFVVSENQS--CFLFMDNI 205
           | |  |: : |     ::| ||                  | | |  |   :    |:
Db       3 HPVYWFLISSSLFASNSLSFANDAQTALTPSDSYNGNVTSEEFQVKETSSGTTYTCEGNV 62

Qy     206 CIQTNTAGKGGAI----YAGTSN-SFESNNCDLFFINNACCA----------------- 242
           ||  : |||    :    :: | | :|  |   | | |     |
Db      63 CI--SFAGKDSGLKKSCFSATDNLTFLGNGYTLCFDNITTTASNPGAINVQGQGKTLGIS 120

Qy     243 GGAIFSPICS-----LTG-----NRGNIVFYNNRCF---KNVETASSEASDGGAIKVTTR 289
           | ::||  |:      ||      :||    :|     ||  ||     :||||:
Db     121 GFSLFS--CAYCPPGTTGYGAIQTKGNTTLKDNSSLVFHKNCSTA-----EGGAIQCKGS 173
```

```
Qy    290 LDV---TGNRGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANN----KGGAIYI 342
Db    174 SDAELKIENNQNLVFSENSSTSKGGAIYADKLTIVSGGPTLFSNNSVSNGSSPKGGAISI 233

Qy    343 -DGTSNSKISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEIL-LGAGS 400
Db    234 KDSSGECSLTADLGDITFDGNKIIKTSGGSSTVT------RNSIDL--GTGKFTKLRAKD 285

Qy    401 SQNLIFYDPIEVSNAGVSVSFNKE--ADQTGSVVFSGATVNSAD-FHQRNLQTKTPAPLT 457
Db    286 GFGIFFYDPI-TGGGSDELNINKKETVDYTGKIVFSGEKLSDEEKARAENLASTFNQPIT 344

Qy    458 LSNGFLCIEDHAQLTVNRFTQ-TGGVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNL 516
Db    345 LSAGSLVLKDGVSVTAKQVTQEAGSTVVMDLGTTLQTPSSG------GETITLTNLDINI 398

Qy    517 SSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALS 576
Db    399 ASLGGGGG------TSPAKLATN-TASQAITIN----AVNLVDADGNA-YEDPILA---T 443

Qy    577 SQPMLSISEASDNQLRSDDMDFSGLNVP--HYGWQGLWTWGWAKTQDPEPASSATITDPQ 634
Db    444 SKPFTAIVATTNASTVTQPTDNLTNYVPPTHYGYQGNWTVTW----DTETAT-------- 491

Qy    635 KANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWG--NMLLATESLKN-SAELTPSDHPF 691
Db    492 ------KTATLTWEQTGYSPNPERQGPLVPNTLWGAFSDLRAIQNLMDISVNGADYHRGF 545

Qy    692 WGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIA--GQTHTFSLKFSQTYTKLNERYA 749
Db    546 W---VSGLANFLHKSGSDTKRKFRHNSAGYALGVYAKTPSDDIFSAAFCQLFGK-----D 597

Qy    750 KNNVSSKNYSCQGEMLFSLQEGFLLTKLVGLYSYGDH------------------NCHH 790
Db    598 KDYLVSKNNA-------NIYAGSLYYQHISYWSAWQNLLQNTIGAEAPLVLNAQLTYCHA 650

Qy    791 FYTQGENLTS------------QGTFRSQTMGGAVFFDLPMKPFGSTHI-LTAPFLGALG 837
Db    651 SNDMKTNMTTTYAPRKTTYAEIKGDWGNDCFGVELGATVPIQTESSLLFDMYSPFLKFQL 710

Qy    838 IYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQ 897
Db    711 VHTHQDDFKENNSDQGRYFESSNLTNLSLPIGIKFERF-ANNDTASYHVTAAYSPDIVRS 769

Qy    898 EPGIATQLLAS--KGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLH-----FQYHGFY-- 948
Db    770 NPDCTTSLLVSPDSAVWVTKANNLARSAF------MLQAGNYLSLSHNIEIFSQFGFELR 823

Qy    949 -SSSTF 953
Db    824 GSSRTY 829


RESULT 14
US-09-841-132-394
; Sequence 394, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
;  APPLICANT: Bhatia, Ajay
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C8
;  CURRENT APPLICATION NUMBER: US/09/841,132
;  CURRENT FILING DATE:  2001-04-23
;  NUMBER OF SEQ ID NOS: 599
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 394
;   LENGTH: 1723
;   TYPE: PRT
;   ORGANISM: Chlamydia pneumoniae

US-09-841-132-394

  Query Match             10.2%;  Score 517;  DB 10;  Length 1723;
  Best Local Similarity   22.8%;  Pred. No. 2.9e-31;
  Matches  257;  Conservative  164;  Mismatches  446;  Indels  258;  Gaps   50;

Qy     15 SGLAREVPSRIFL------MPNSVPDPTKESLSNKISLTGDTHNLT-------NCYLDNL 61
Db    683 NGGGANVPENIVLTFTYTPTPNE-PAPVQQPVYGEALVTGNTATKSGGGIYTKNAAFSNL 741

Qy     62 RYILAILQKTPNEGAAVTIT----DYLSFFDTQKEGIYFAKNLTPESGGAIG-----YAS 112
Db    742 SSV-TFDQNTSSENGGALLTQKAADKTDCSFTYITNVNITNNTATGNGGGIAGGKAHFDR 800

Qy    113 PNSPTVEIRDTI--GPVIFENNTCCRPFT--SSNPNAAVNKIREGGAIHAQNLYIN---- 164
Db    801 IDNLTVQSNQAKKGGGVYLEDALILEKVITGSVSQNTAT---ESGGGIYAKDIQLQALPG 857

Qy    165 ----HNHDVVGFMKNFSYVRGGAISTANTFVVSENQSCFLFMDNICIQTNTAG----KGG 216
Db    858 SFTITDNKVETSLTTSTNLYGGGIYSSGAVTLTNISGTFGITGNSVINTATSQDADIQGG 917

Qy    217 AIYAGTSNSFESNNCDLFFINNACC----------AGGAIFSPICSLTGNRGNIVFYNNR 266
Db    918 GIYATTSLSINQCNTPILFSNNSAATKKTSTTKQIAGGAIFSAAVTIENNSQPIIFLNNS 977

Qy    267 CFKNVETAS---SEASDGGAI----------KVTTR-----------LDVT------- 293
Db    978 AKSEATTAATAGNKDSCGGAIAANSVTLTNNPEITFKGNYAETGGAIGCIDLTNGSPPRK 1037

Qy    294 ---GNRGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKI 350
Db   1038 VSIADNGSVLFQDNSALNRGGAIYGETIDISRTGAT-FIGNSSKHDGSAIC---CSTALT 1093

Qy    351 SADRHAIIFNENIVTNVT--------NANGTSTSANPPRRNAITVASSSGEIL------- 395
Db   1094 LAPNSQLIFENNKVTETTATTKASINNLGAAIYGNNETSDVTISLSAENGSIFFKNNLCT 1153

Qy    396 ----------------LGAGSSQNLIFYDPIEVSNAGVS---VSFNKEADQTGSVVFSGA 436
Db   1154 ATNKYCSIAGNVKFTAIEASAGKAISFYDAVNVSTKETNAQELKLNEKATSTGTILFSG- 1212

Qy    437 TVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQT-GGVVSLGNGAVLSCYK 495
Db   1213 -----ELHEN--KSYIPQKVTFAHGNLILGKNAELSVVSFTQSPGTTITMGPGSVLSNHS 1265

Qy    496 NGAGNSASNASITLKHIGLNLSSIL--KSGAEI---PLLWVEPTNNSNNYTADTAATFSL 550
Db   1266 KEAG------GIAINNVIIDFSEIVPTKDNATVAPPTLKLVSRTNADSKDKIDITGTVTL 1319

Qy    551 SDVKLSLIDDYGNSPY-ESTDLTHALSSQPMLSISEASDNQLRSDDMDFSGLNVPHYGWQ 609
Db   1320 LDPNGNL---YQNSYLGEDRDIT-------LFNIDNSASGAVTATNVTLQGNLGAKKGYL 1369

Qy    610 GLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTW-----LPAGYVPSPKHRSPLIA 664
Db   1370 GTW------NLDPNSSGS--------------KIILKWTFDKYLRWPYIPRDNH---FYI 1406

Qy    665 NTLWG--NMLLATES------LKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHM 716
Db   1407 NSIWGAQNSLVTVKQGILGNMLNNARFEDPAFNNFW---ASAIGSFLRKEVSRNSDSFTY 1463

Qy    717 RSSGYSAGMIAGQTHTFSL--KFSQTYTKLNERYAKNNVSSK--NYSCQ-----GEMLF- 766
Db   1464 HGRGYTAAVDAKPRQEFILGAAFSQVFGHAESEYHLDNYKHKGSGHSTQASLYAGNIFYF 1523

Qy    767 -SLQEGFLLTKLVGLYSYGDHNCHHFY--TQGENLTSQGTFRSQTMGGAVFFDLPM---- 819
Db   1524 PAIRSRPILFQGVATYGYMQHDTTTYYPSIEEKNMANWDSI-------AWLFDLRFSVDL 1576

Qy    820 ---KPFGSTHILTAPFLGALGIYSSLSH-------FTEVGAYPRSFSTKTPLINVLVPIG 869
Db   1577 KEPQPHSTAR---------LTFYTEAEYTRIRQEKFTELDYDPRSFSA-CSYGNLAIPTG 1626
```

```
Qy    870 --VKGSFMNATQRPQAW-------TVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSS 920
          | |:         ||        |  || ||: |  |    ::|::|          :
Db   1627 FSVDGAL--------AWREIILYNKVSAAYLPVILRNNPKATYEVLSTKEKGNVVNVLPT 1678

Qy    921 RHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
          |:|   ::| |    |: ||:  |    | :|     || |   |
Db   1679 RNAARAEVSSQIYLGSYWTLYGTYTIDASMNTLVQMANGGIRFVF 1723


RESULT 15
US-09-841-132-395
; Sequence 395, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
;  APPLICANT: Bhatia, Ajay
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C8
;  CURRENT APPLICATION NUMBER: US/09/841,132
;  CURRENT FILING DATE:  2001-04-23
;  NUMBER OF SEQ ID NOS: 599
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 395
;   LENGTH: 1723
;   TYPE: PRT
;   ORGANISM: Chlamydia pneumoniae
US-09-841-132-395

  Query Match             10.2%;  Score 517;  DB 10;  Length 1723;
  Best Local Similarity   22.8%;  Pred. No. 2.9e-31;
  Matches  257;  Conservative  164;  Mismatches  446;  Indels  258;  Gaps   50;

Qy     15 SGLAREVPSRIFL------MPNSVPDPTKESLSNKISLTGDTHNLT-------NCYLDNL 61
          :|    ||  | |       ||  | | :: :  :  :||:|   :       |    ||
Db    683 NGGGANVPENIVLTFTYTPTPNE-PAPVQQPVYGEALVTGNTATKSGGGIYTKNAAFSNL 741

Qy     62 RYILAILQKTPNEGAAVTIT----DYLSFFDTQKEGIYFAKNLTPESGGAIG-----YAS 112
            :   | | :|     :|    |      |    :    |    :|| |      :
Db    742 SSV-TFDQNTSSENGGALLTQKAADKTDCSFTYITNVNITNNTATGNGGGIAGGKAHFDR 800

Qy    113 PNSPTVEIRDTI--GPVIFENNTCCRPFT--SSNPNAAVNKIREGGAIHAQNLYIN---- 164
           :: ||:       | |  |:          | : | |      || |:|::: :
Db    801 IDNLTVQSNQAKKGGGVYLEDALILEKVITGSVSQNTAT---ESGGGIYAKDIQLQALPG 857

Qy    165 ----HNHDVVGFMKNFSYVRGGAISTANTFVVSENQSCFLFMDNICIQTNTAG----KGG 216
              :: |   :   : : || | ::    ::     |    |  | | |:      :||
Db    858 SFTITDNKVETSLTTSTNLYGGGIYSSGAVTLTNISGTFGITGNSVINTATSQDADIQGG 917

Qy    217 AIYAGTSNSFESNNCDLFFINNACC----------AGGAIFSPICSLTGNRGNIVFYNNR 266
           ||| || |    |  : | ||:           |||||||   ::  |   |:| ||
Db    918 GIYATTSLSINQCNTPILFSNNSAATKKTSTTKQIAGGAIFSAAVTIENNSQPIIFLNNS 977

Qy    267 CFKNVETAS---SEASDGGAI----------KVTTR-----------LDVT------- 293
              ||:   :: | ||||           ::| :            :|:|
Db    978 AKSEATTAATAGNKDSCGGAIAANSVTLTNNPEITFKGNYAETGGAIGCIDLTNGSPPRK 1037

Qy    294 ---GNRGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKI 350
             : | : | ||   | |||||   : :   | | || | : : | ||    | :
Db   1038 VSIADNGSVLFQDNSALNRGGAIYGETIDISRTGAT-FIGNSSKHDGSAIC---CSTALT 1093

Qy    351 SADRHAIIFNENIVTNVT--------NANGTSTSANPPRRNAITVASSSGEIL------- 395
          |    :||  | ||  |          |         |      |:::: :| |
Db   1094 LAPNSQLIFENNKVTETTATTKASINNLGAAIYGNNETSDITISLSAENGSIFFKNNLCT 1153

Qy    396 ----------------LGAGSSQNLIFYDPIEVSNAGVS---VSFNKEADQTGSVVFSGA 436
                          : | : : : ||| : ||    :   :  |::|  ||:::|||
Db   1154 ATNKYCSIAGNVKFTAIEASAGKAISFYDAVNVSTKETNAQELKLNEKATSTGTILFSG- 1212

Qy    437 TVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQT-GGVVSLGNGAVLSCYK 495
             : |:   ::  |  :| ::| | :  :|:|:|  |||: |  :::| |:||| :
```

```
Db   1213 -----ELHEN--KSYIPQKVTFAHGNLILGKNAELSVVSFTQSPGTTITMGPGSVLSNHS 1265

Qy    496 NGAGNSASNASITLKHIGLNLSSIL--KSGAEI---PLLWVEPTNNSNNYTADTAATFSL 550
            ||       | : :: :: | |:  |  | :    |  |  ||  :    |   | :|
Db   1266 KEAG------GIAINNVIIDFSEIVPTKDNATVAPPTLKLVSRTNADSKDKIDITGTVTL 1319

Qy    551 SDVKLSLIDDYGNSPY-ESTDLTHALSSQPMLSISEASDNQLRSDDMDFSGLNVPHYGWQ 609
           |   :|   | ||   |  |:|       : :|  ::   : : ::   |      |:
Db   1320 LDPNGNL---YQNSYLGEDRDIT-------LFNIDNSASGAVTATNVTLQGNLGAKKGYL 1369

Qy    610 GLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTW-----LPAGYVPSPKHRSPLIA 664
          | |        ||  : |                ::| |     |   |:|   |
Db   1370 GTW------NLDPNSSGS-------------KIILKWTFDKYLRWPYIPRDNH---FYI 1406

Qy    665 NTLWG--NMLLATES------LKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHM 716
          |::||  | |:  :       | |:    |: : ||      :|  : ::   |   |
Db   1407 NSIWGAQNSLVTVKQGILGNMLNNARFEDPAFNNFW---ASAIGSFLRKEVSRNSDSFTY 1463

Qy    717 RSSGYSAGMIAGQTHTFSL--KFSQTYTKLNERYAKNNVSSK--NYSCQ-----GEMLF- 766
             ||:| : |      | |   |||  :       |  :|   |   :| | |     | : :
Db   1464 HGRGYTAAVDAKPRQEFILGAAFSQVFGHAESEYHLDNYKHKGSGHSTQASLYAGNIFYF 1523

Qy    767 -SLQEGFLLTKLVGLYSYGDHNCHHFY--TQGENLTSQGTFRSQTMGGAVFFDLPM---- 819
           :::   :| : |  | |  |:   :|   : :|: :  :         |  |||
Db   1524 PAIRSRPILFQGVATYGYMQHDTTTYYPSIEEKNMANWDSI-------AWLFDLRFSVDL 1576

Qy    820 ---KPFGSTHILTAPFLGALGIYSSLSH-------FTEVGAYPRSFSTKTPLINVLVPIG 869
             :|  :            |  |:   :      |||:   |||||      |: :| |
Db   1577 KEPQPHSTAR---------LTFYTEAEYTRIRQEKFTELDYDPRSFSA-CSYGNLAIPTG 1626

Qy    870 --VKGSFMNATQRPQAW-------TVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSS 920
          | |:         ||        |  || ||: |  |    ::|::|          :
Db   1627 FSVDGAL--------AWREIILYNKVSAAYLPVILRNNPKATYEVLSTKEKGNVVNVLPT 1678

Qy    921 RHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
          |:|   ::| |    |: ||:  |    | :|     || |   |
Db   1679 RNAARAEVSSQIYLGSYWTLYGTYTIDASMNTLVQMANGGIRFVF 1723


Search completed: February  4, 2003, 21:08:42
Job time : 81 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2003, 20:53:42 ; Search time 94 Seconds
(without alignments)
2115.271 Million cell updates/sec

Title: US-09-677-752-2
Perfect score: 5086
Sequence: 1 MKKAFFFFLIGNSLSGLARE.....GFYSSSTFCNYLNGEIALRF 965

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	636	12.5	926	2	P71135	P71135 chlamydophi
2	544.5	10.7	847	2	P71132	P71132 chlamydophi
3	531	10.4	839	2	P77792	P77792 chlamydia p
4	530.5	10.4	772	16	Q9RB71	Q9rb71 chlamydia p
5	530.5	10.4	846	2	P71133	P71133 chlamydophi
6	512	10.1	392	16	Q9Z881	Q9z881 chlamydia p
7	477	9.4	700	2	Q8VU50	Q8vu50 chlamydia p
8	422	8.3	347	16	Q9RB61	Q9rb61 chlamydia p
9	387	7.6	1530	2	Q93QE8	Q93qe8 chlamydia t
10	349	6.9	602	2	Q8VU49	Q8vu49 chlamydia p
11	346.5	6.8	581	2	Q8VU48	Q8vu48 chlamydia p
12	344.5	6.8	601	2	Q8VL57	Q8vl57 chlamydia p
13	321	6.3	494	16	Q9RB68	Q9rb68 chlamydia p
14	297	5.8	427	16	Q9RB70	Q9rb70 chlamydia p
15	275.5	5.4	445	16	Q9RB67	Q9rb67 chlamydia p
16	253.5	5.0	359	16	Q9JSK6	Q9jsk6 chlamydia p
17	249	4.9	649	2	P71134	P71134 chlamydophi
18	234.5	4.6	1250	16	Q8XE28	Q8xe28 escherichia
19	224	4.4	989	2	Q9XD84	Q9xd84 escherichia
20	213	4.2	2902	16	Q9ZME6	Q9zme6 helicobacte
21	206	4.1	1364	2	Q8VSL2	Q8vsl2 shigella fl
22	205.5	4.0	820	12	Q89364	Q89364 paramecium
23	204	4.0	1604	2	Q9KK99	Q9kk99 rickettsia
24	203	4.0	1654	2	Q93QW9	Q93qw9 rickettsia
25	202	4.0	2399	16	Q9ZKS9	Q9zks9 helicobacte
26	201.5	4.0	1487	16	Q8YK40	Q8yk40 anabaena sp
27	201.5	4.0	1881	16	Q8RGK2	Q8rgk2 fusobacteri
28	201	4.0	1029	2	Q52708	Q52708 rickettsia
29	200.5	3.9	1369	12	Q89349	Q89349 paramecium
30	200.5	3.9	1620	2	Q9KKB5	Q9kkb5 rickettsia
31	200	3.9	2893	16	O25063	O25063 helicobacte
32	199.5	3.9	1102	2	Q52298	Q52298 shigella fl
33	199.5	3.9	1102	2	Q99Q93	Q99q93 shigella fl
34	198.5	3.9	1335	12	Q89353	Q89353 paramecium
35	197.5	3.9	1304	2	Q9XC46	Q9xc46 rickettsia
36	196.5	3.9	1649	16	Q9CFA2	Q9cfa2 lactococcus
37	195	3.8	1364	2	Q99QC6	Q99qc6 shigella fl
38	195	3.8	1618	2	Q9KKB4	Q9kkb4 rickettsia
39	195	3.8	1619	2	Q9KKB7	Q9kkb7 rickettsia
40	194	3.8	1603	2	Q9KKA6	Q9kka6 rickettsia
41	194	3.8	1643	2	Q9F0P7	Q9f0p7 rickettsia
42	193	3.8	1616	2	Q9KKA5	Q9kka5 rickettsia
43	192.5	3.8	2529	16	O25579	O25579 helicobacte
44	192	3.8	1070	16	Q8ZA36	Q8za36 yersinia pe
45	192	3.8	1072	2	Q9F288	Q9f288 yersinia pe

ALIGNMENTS

RESULT 1
P71135
ID P71135 PRELIMINARY; PRT; 926 AA.
AC P71135;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative 98 kDa outer membrane protein.
OS Chlamydophila abortus.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83555;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OVINE ABORTION S26/3;
RA Longbottom D., Russell M., Dunbar S.M., Jones G.E., Herring A.J.;
RT "98kDa protein genes from ovine abortion strain S26/3 Chlamydia
RT psittaci.";
RL Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U72499; AAB18188.1; -.
DR InterPro; IPR003368; Chlamydia_PMP.
DR InterPro; IPR003357; OMP.
DR Pfam; PF02415; DUF145; 2.
DR Pfam; PF02385; OMP; 1.
SQ SEQUENCE 926 AA; 98439 MW; 3E755E52F594750F CRC64;

Query Match 12.5%; Score 636; DB 2; Length 926;
Best Local Similarity 24.0%; Pred. No. 5.4e-29;
Matches 241; Conservative 157; Mismatches 365; Indels 240; Gaps 41;

```
Qy      41 SNKISLTGDTHNL--TNCYLDNLRYILAILQKTPNEGA-----------AVTITDY--L 84
           :: ::  |: |:|  ||                | ||             :|:||:  |
Db      82 ADNLTFKGNNHSLSITNA---------------NAGANPAGINVNTADKILTLTDFSKL 125

Qy      85 SFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNPN 144
           || :    :  |   :||||:  |:  |          ::|: |    ::: |
Db     126 SFKECPSSLVNTGKG-AMKSGGALNLANNAS----------ILFDQN-----YSAEN-- 166

Qy     145 AAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSCFLFMDN 204
                ||||  :   :  :   : |  | :  :||||:      :|:||    |  |
```

```
Db    167 --------GGAISCKAFSLTGSSKEISFTTNSTAKKGGAIAATGIAHLSDNQGTIRFSGN 218

Qy    205 ICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNACC-----AGGAIFSPICSLTGNRGN 259
Db    219 TAVNS-----GGAVYSEASMTIAGNN-HVAFSNNAVSGSSDGCGGAIH---CSKTGSAPT 269

Qy    260 IVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYAPVVT 319
Db    270 LTIRDNKV----------------------------LIFEENTSSAKGGAIYTDKLI 298

Qy    320 LVDNGPTYFINN---IANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTSTS 376
Db    299 LTSGGPTAFINNKVTHATPKGGAIGIAANGECSLTAEHGDITFDNNLMATQDNAT----- 353

Qy    377 ANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEV-SNAGVSVSFNK-EADQT--GSVV 432
Db    354 ---IKRNAINIEGNGKFVNLRAASGKTISFYDPITVEGNAADLLTLNKAEGDKTYNGRII 410

Qy    433 FSGATVNS-----ADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGN 487
Db    411 FSGEKLTEEQAAVAD----NLKTTFTQPITLAAGELVLRSGVEVEAKTVVQTAGSLILMD 466

Qy    488 GAVLSCYKNGAGNSASNASITLKHIGLNLSSI----------LKSGAEIPLLW----VEP 533
Db    467 A--------GTKLSAKTEDATLTNLAINPNTLDGKKFAVVDAVAAGKNVTLSGAIGVIDP 518

Qy    534 TNN--SNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDL-THALSSQPMLSISEASDNQ 590
Db    519 TGKFYENHKLNDTLA---LGGIQLS-----GKGSVTTTNVPSHVV--------------- 555

Qy    591 LRSDDMDFSGLNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPA 650
Db    556 ---------GVAETHYGYQGNWSVSWVKDNNSDPKT-------------QTAIFTWNKT 592

Qy    651 GYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSD--HPFWGITGGGLGMMVYQDPR 708
Db    593 GYVPNPERRAPLVLNSLWGS-FIDLRSIQDVLERSVDSILETRRGLWVSGIGNFFHKDRN 651

Qy    709 ENHPGFHMRSSGYSAGMIAGQTH--TFSLKFSQTYTKLNERYAKNNVSS---------- 755
Db    652 AENRKFRHISSGYVLGATTNTSREDSLSVAFCQLFAKDKDYLVSKNAANVYAGSVYYQHV 711

Qy    756 ------------KNYSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQGENLTSQGT 803
Db    712 SKFDDLTRLFNGPNTCCSG---FSKEIPIFLDAQI-TYCHTANNMTTSYTDYPEV--KGS 765

Qy    804 FRSQTMGGAVFFDLPMKPFGSTHILT-APFLGALGIYSSLSHFTEVGAYPRSFSTKTPLI 862
Db    766 WGNDTLGLTLSTSVPIPVFSSSIFDSYAPFAKLQVVYAHQDDFKEPTTEGRVFES-SDLL 824

Qy    863 NVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRH 922
Db    825 NVSVPIGIKFEKLSYGER-SAYDLTLMYIPDVYRHNPSCMTGLAINDVSWLTTATNLARQ 883

Qy    923 AMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNY---LNGEIA 962
Db    884 AFIVRAGNHIALTSGVEMFSQF-GFELRSSSRNYNVDLGAKVA 925
```

```
RESULT 2
P71132
ID   P71132      PRELIMINARY;      PRT;   847 AA.
AC   P71132;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   POMP91A.
OS   Chlamydophila abortus.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83555;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S26/3;
RX   MEDLINE=96406378; PubMed=8810511;
RA   Longbottom D., Russell M., Jones G.E., Lainson F.A., Herring A.J.;
RT   "Identification of a multigene family coding for the 90 kDa proteins
RT   of the ovine abortion subtype of Chlamydia psittaci.";
RL   FEMS Microbiol. Lett. 142:277-281(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S26/3;
RX   MEDLINE=98187897; PubMed=9529048;
RA   Longbottom D., Russell M., Dunbar S.M., Jones G.E., Herring A.J.;
RT   "Molecular cloning and characterization of the genes coding for the
RT   highly immunogenic cluster of 90-kilodalton envelope proteins from the
RT   Chlamydia psittaci subtype that causes abortion in sheep.";
RL   Infect. Immun. 66:1317-1324(1998).
DR   EMBL; U65942; AAC15921.1; -.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   InterPro; IPR003357; OMP.
DR   Pfam; PF02415; DUF145; 1.
DR   Pfam; PF02385; OMP; 1.
SQ   SEQUENCE   847 AA;  90694 MW;  754C958E7F11179E CRC64;
```

Query Match 10.7%; Score 544.5; DB 2; Length 847;
Best Local Similarity 25.6%; Pred. No. 1.1e-23;
Matches 222; Conservative 126; Mismatches 349; Indels 171; Gaps 43;

```
Qy    129 FENNTCCRPFTSSNPNAAVNKIREGGAIH--AQNLYINHNHDVVGFMKNFSYVRGGAIST 186
Db     36 YNGNTNSEPF---NPLSTSN---SNGTIYTCTGNICI----------AYAGLDGSGLSS 78

Qy    187 ANTFVVSENQSCFL-------FMDN---IC---IQTNTAGKGGAIYAGTSNSFESNNCDL 233
Db     79 ----------SCFTDTAGNLSFLGNGYTLCFDNITTQSSHPGAISVSGTNKTLDISGFSL 128

Qy    234 FFINNACCAGGAIFSPICSLTGN-----RGNIVFYNNRCFKNVETASSEASDGGAIKV-- 286
Db    129 F--SCAYCPPGATGYGAIKAVGNTTIKDNSSLVFHKN----------CSTGEGGAIQCKA 176

Qy    287 ---TTRLDVTGNRGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINN---IANNKGGAI 340
Db    177 SSSEAELKIENNQ-NLVFAENSSSSSGGAIYADKLTIVSGGPTLFSNNSVSASSPKGGAI 235

Qy    341 YI-DGTSNSKISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAG 399
Db    236 CIKDSGGECSLTADLGDITFDGN---KIIKTNGGSPTVT---RNSIDLGSSGKFTKLNAK 289

Qy    400 SSQNLIFYDPIEVSNAGVSVSFNKE--ADQTGSVVFSGATVNSADFH-QRNLQTKTPAPL 456
Db    290 EGFGIFFYDPI-TGGGSDELNINKQDTVDYTGKIVFSGERLSDEEKKVAANLKSDFKQPL 348

Qy    457 TLSNGFLCIEDHAQLTVNRFTQT-GGVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLN 515
Db    349 KIGSGSLILKDGVTLETKSFTQTEGATVVMDLGTTLQTPSSG------GETITLTNLDIN 402

Qy    516 LSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTHAL 575
Db    403 VASLGGGGVAPDPAKVEATTESKTVTINA----------VNLVDDNGNA-YEYPILA--- 448

Qy    576 SSQPM--LSISEASDNQLRSDDMDFSGLNVP-HYGWQGLWTWGWAKTQDPEPASSATITD 632
Db    449 ASQPFTAIEVRSGSSGSITKPTTNLENYTPPTHYGYQGNWTVTW------KQGSSAQ--- 499

Qy    633 PQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNM--LLATESLKN-SAELTPSDH 689
Db    500 -------EKTATLTWEQTGYSPNPERQGSLVPNTLWGSFSDIRAIQNLMDISVNGADYHR 552

Qy    690 PFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQT---HTFSLKFSQTYTKLNE 746
Db    553 GFW---VSGLGNFLHKSGSDTKRKFRHNSAGYALGVYA-QTPSEDVFSAAFCQLFGK--- 605

Qy    747 RYAKNNVSSKNYSC--QGEMLFSLQE-----GFLLTKLVGL---------YSYGDHNCHH 790
Db    606 --DKDYLVSKNSSTVYAGSIYYQHISYWNTWNTLLQNTLGAEAPLVLNAQLAY----CHA 659
```

```
Qy    791 FYTQGENLTS------------QGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGI 838
Db    660 SNNMKTNMTDTYAPPKTTYSEIKGDWGNDCFGVEFGAKAPIETASLLFDMYSPFVKLQLV 719

Qy    839 YSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVK-GSFMNATQRPQAWTVELAYQPVLYRQ 897
Db    720 HAHQDDFKENNSDQGRYFESNNLTNLSMPIGVKLEKFSH--KDTASYNLTLAYAPDIVRS 777

Qy    898 EPGIATQLLAS--KGIWFGSGSPSSRHA 923
Db    778 NPDCTASLLVSPTSAVWVTKANNLARHA 805


RESULT 3
P77792
ID   P77792      PRELIMINARY;      PRT;   839 AA.
AC   P77792;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   POMP90B precursor.
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OVINE ENZOOTIC ABORTION ISOLATE S26/3;
RA   Longbottom D., Russell M., Jones G.E., Lainson F.A., Herring A.J.;
RL   Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OVINE ENZOOTIC ABORTION ISOLATE S26/3;
RA   Longbottom D., Russell M., Dunbar S.M., Jones G.E., Herring A.J.;
RL   Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U65943; AAC15924.1; -.
DR   EMBL; U65942; AAC15922.1; -.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   InterPro; IPR003357; OMP.
DR   Pfam; PF02415; DUF145; 1.
DR   Pfam; PF02385; OMP; 1.
KW   Signal.
FT   SIGNAL        1     16       POTENTIAL.
FT   CHAIN        17    839       POMP90B.
SQ   SEQUENCE   839 AA;  89824 MW;  4581C7CBAF7FF4C4 CRC64;

  Query Match             10.4%;  Score 531;  DB 2;  Length 839;
  Best Local Similarity   25.6%;  Pred. No. 6.6e-23;
  Matches  232;  Conservative  129;  Mismatches  347;  Indels  198;  Gaps   44;

Qy    167 HDVVGFMKNFSYVRGGAISTAN------------------TFVVSENQS--CFLFMDNI 205
Db      3 HPVYWFLISSSLFASNSLSFANDAQTALTPSDSYNGNVTSEEFQVKETSSGTTYTCEGNV 62

Qy    206 CIQTNTAGKGGAI----YAGTSN-SFESNNCDLFFINNACCA----------------- 242
Db     63 CI--SFAGKDSGLKKSCFSATDNLTFLGNGYTLCFDNITTTASNPGAINVQGQGKTLGIS 120

Qy    243 GGAIFSPICS-----LTG-----NRGNIVFYNNRCF---KNVETASSEASDGGAIKVTTR 289
Db    121 GFSLFS--CAYCPPGTTGYGAIQTKGNTTLKDNSSLVFHKNCSTA-----EGGAIQCKGS 173

Qy    290 LDV---TGNRGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANN----KGGAIYI 342
Db    174 SDAELKIENNQNLVFSENSSTSKGGAIYADKLTIVSGGPTLFSNNSVSNGSSPKGGAISI 233

Qy    343 -DGTSNSKISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEIL-LGAGS 400
Db    234 KDSSGECSLTADLGDITFDGNKIIKTSGGSSTVT------RNSIDL--GTGKFTKLRAKD 285

Qy    401 SQNLIFYDPIEVSNAGVSVSFNKE--ADQTGSVVFSGATVNSAD-FHQRNLQTKTPAPLT 457
Db    286 GFGIFFYDPI-TGGGSDELNINKKETVDYTGKIVFSGEKLSDEEKARAENLASTFNQPIT 344

Qy    458 LSNGFLCIEDHAQLTVNRFTQ-TGGVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNL 516
Db    345 LSAGSLVLKDGVSVTAKQVTQEAGSTVVMDLGTTLQTPSSG------GETITLTNLDINI 398

Qy    517 SSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALS 576
Db    399 ASLGGGGG------TSPAKLATN-TASQAITIN----AVNLVDADGNA-YEDPILA---T 443

Qy    577 SQPMLSISEASDNQLRSDDMDFSGLNVP--HYGWQGLWTWGWAKTQDPEPASSATITDPQ 634
Db    444 SKPFTAIVATTNASTVTQPTDNLTNYVPPTHYGYQGNWTVTW----DTETAT-------- 491

Qy    635 KANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWG--NMLLATESLKN-SAELTPSDHPF 691
Db    492 ------KTATLTWEQTGYSPNPERQGPLVPNTLWGAFSDLRAIQNLMDISVNGADYHRGF 545

Qy    692 WGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIA--GQTHTFSLKFSQTYTKLNERYA 749
Db    546 W---VSGLANFLHKSGSDTKRKFRHNSAGYALGVYAKTPSDDIFSAAFCQLFGK-----D 597

Qy    750 KNNVSSKNYSCQGEMLFSLQEGFLLTKLVGLYSYGDH------------------NCHH 790
Db    598 KDYLVSKNNA-------NIYAGSLYYQHISYWSAWQNLLQNTIGAEAPLVLNAQLTYCHA 650

Qy    791 FYTQGENLTS------------QGTFRSQTMGGAVFFDLPMKPFGSTHI-LTAPFLGALG 837
Db    651 SNDMKTNMTTTYAPRKTTYAEIKGDWGNDCFGVELGATVPIQTESSLLFDMYSPFLKFQL 710

Qy    838 IYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQ 897
Db    711 VHTHQDDFKENNSDQGRYFESSNLTNLSLPIGIKFERF-ANNDTASYHVTAAYSPDIVRS 769

Qy    898 EPGIATQLLAS--KGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLH-----FQYHGFY-- 948
Db    770 NPDCTTSLLVSPDSAVWVTKANNLARSAF------MLQAGNYLSLSHNIEIFSQFGFELR 823

Qy    949 -SSSTF 953
Db    824 GSSRTY 829


RESULT 4
Q9RB71
ID   Q9RB71      PRELIMINARY;      PRT;   772 AA.
AC   Q9RB71;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Pmp_3.
GN   PMP_3_2.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
DR   EMBL; AP002545; BAA98226.1; -.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   InterPro; IPR003357; OMP.
DR   Pfam; PF02415; DUF145; 1.
DR   Pfam; PF02385; OMP; 1.
SQ   SEQUENCE   772 AA;  82931 MW;  484FC56D635801EB CRC64;

  Query Match             10.4%;  Score 530.5;  DB 16;  Length 772;
  Best Local Similarity   25.8%;  Pred. No. 6.3e-23;
  Matches  221;  Conservative  124;  Mismatches  343;  Indels  167;  Gaps   39;
```

```
Qy    172 FMKNFSYVRGGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNC 231
Db      3 FSKNFSTDNGGAI-TAKTLSLTGTTMSALFSEN-----TSSKKGGAIQTSDALTITGNQG 56

Qy    232 DLFFINNACC-AGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEAS---DGGAI--- 284
Db     57 EVSFSDNTSSDSGAAIFTEASVTISNNAKVSFIDNK----VTGASSSTTGDMSGGAICAY 112

Qy    285 --KVTTRLDVTGNRGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANN----KGG 338
Db    113 KTSTDTKVTLTGNQ-MLLFSNNTSTTAGGAIYVKKLELASGGLTLFSRNSVNGGTAPKGG 171

Qy    339 AIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGA 398
Db    172 AIAIEDSGELSLSADSGDIVFLGNTVTSTT--PGTN-------RSSIDLGTSAKMTALRS 222

Qy    399 GSSQNLIFYDPIEVSNAGVSVSFNK----EADQ----TGSVVFSGATVNSAD-FHQRNLQ 449
Db    223 AAGRAIYFYDPITTGSSTTVTDVLKVNETPADSALQYTGNIIFTGEKLSETEAADSKNLT 282

Qy    450 TKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNSASNA-SIT 508
Db    283 SKLLQPVTLSGGTLSLKHGVTLQTQAFTQQAD----------SRLEMDVGTTLEPADTST 332

Qy    509 LKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYES 568
Db    333 INNLVINISSI--DGAK--KAKIETKATSKNLT--LSGTITLLDPTGTFYENHSLRNPQS 386

Qy    569 TDLTHALSSQPMLSISEASDNQLRSDDMDFSGLNVP------HYGWQGLW---TWGWAKT 619
Db    387 YDILELKASGTVTSTAVTPD--------------PIMGEKFHYGYQGTWGPIVWG---- 427

Qy    620 QDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESL- 678
Db    428 ----TGASTTAT-------------FNWTKTGYIPNPERIGSLVPNSLW-NAFIDISSLH 469

Qy    679 ---KNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFSL 735
Db    470 YLMETANEGLQGDRAFW---CAGLSNFFHKDSTKTRRGFRHLSGGY---VIGGNLHTCSD 523

Qy    736 K-----FSQTYTKLNERY-AKNN---VSSKNYSCQGEMLFSL-------QEGFLLTKLVG 779
Db    524 KILSAAFCQLFGRDRDYFVAKNQGTVYGGTLYYQHNETYISLPCKLRPCSLSYVPTEIPV 583

Qy    780 LYSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILT---------- 829
Db    584 LFS---GNLSYTHTDNDLKTKYTTY--PTVKGSWGNDSFALEFGGRAPICLDESALFEQY 638

Qy    830 APFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVK----GSFMNATQRPQAWT 885
Db    639 MPFMKLQFVYAHQEGFKEQGTEAREFGS-SRLVNLALPIGIRFDKESDCQDAT-----YN 692

Qy    886 VELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYH 945
Db    693 LTLGYTVDLVRSNPDCTTTLRISGDSWKTFGTNLARQALVLRAGN----------HFCFN 742

Qy    946 -GFYSSSTFCNYLNG 959
Db    743 SNFEAFSQFSFELRG 757


RESULT 5
P71133
ID   P71133      PRELIMINARY;      PRT;   846 AA.
AC   P71133;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   POMP91B precursor.
OS   Chlamydophila abortus.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83555;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S26/3;
RX   MEDLINE=96406378; PubMed=8810511;
RA   Longbottom D., Russell M., Jones G.E., Lainson F.A., Herring A.J.;
RT   "Identification of a multigene family coding for the 90 kDa proteins
RT   of the ovine abortion subtype of Chlamydia psittaci.";
RL   FEMS Microbiol. Lett. 142:277-281(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S26/3;
RX   MEDLINE=98187897; PubMed=9529048;
RA   Longbottom D., Russell M., Dunbar S.M., Jones G.E., Herring A.J.;
RT   "Molecular cloning and characterization of the genes coding for the
RT   highly immunogenic cluster of 90-kilodalton envelope proteins from the
RT   Chlamydia psittaci subtype that causes abortion in sheep.";
RL   Infect. Immun. 66:1317-1324(1998).
DR   EMBL; U65943; AAC15923.1; -.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   InterPro; IPR003357; OMP.
DR   Pfam; PF02415; DUF145; 1.
DR   Pfam; PF02385; OMP; 1.
KW   Signal.
FT   SIGNAL       1     16       POTENTIAL.
FT   CHAIN       17    846       POMP91B.
SQ   SEQUENCE   846 AA;  90834 MW;  4CDC31DC03C2964E CRC64;

  Query Match             10.4%;  Score 530.5;  DB 2;  Length 846;
  Best Local Similarity   24.2%;  Pred. No. 7.1e-23;
  Matches  239;  Conservative  141;  Mismatches  361;  Indels  245;  Gaps   47;

Qy      1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSV--PDPTKESLSNKISLTGDTHNLTNCYL 58
Db      1 MKHPVYWFLISSSL------------LASNSLSFAQVTNETLTSSDSYNG---NVT---- 41

Qy     59 DNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTV 118
Db     42 ----------------------SDEFEVKETTSGAIY-----TCEGNVCISYAGKDSPL- 73

Qy    119 EIRDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSY 178
Db     74 ------------NKSCFSETT---------------------------------ENLSF 87

Qy    179 VRGGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINN 238
Db     88 IGNG------------YTLCF---DN--ITTQSSHPGAISVSGTNKTLDISGFSLFSCAY 130

Qy    239 ACCAG----GAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKV-----TTR 289
Db    131 CCPPGTTGYGAIQTKGTTTLKDNSSLVFH-----KNCSTA-----EGGAIQCKSSSSTAE 180

Qy    290 LDVTGNRGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINN-IANN---KGGAIYI-DG 344
Db    181 LKLENNK-NLVFSENSSKEKGGAIYADKLTIVSGGPTLFSNNSVSHNSSPKGGAICIKDS 239

Qy    345 TSNSKISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNL 404
Db    240 DGECSLTANLGDITFDGN---KIITTNGGSPTVT---RNSIDLGSGGKFTKLNAKEGFGI 293

Qy    405 IFYDPIEVSNAGVSVSFNK-EAD--QTGSVVFSGATVNSAD-FHQRNLQTKTPAPLTLSN 460
Db    294 FFYDPIANTGGSTEIELNKTESDTTYTGKIVFSGEKLSDEEKTVPANLKSYFKQPLKIGA 353

Qy    461 GFLCIEDHAQLTVNRFTQT-GGVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSI 519
Db    354 GSLVLKDGVTLEAKKITQTKGSTVVMDLGTTLQT------PSSSGETITLTNLDINIASL 407

Qy    520 LKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQP 579
Db    408 GGGGGTAPA-----------KLATNTASQAISIAAVNLVNTDSNT-YE----------DP 445

Qy    580 MLSISEA--------SDNQLRSDDMDFSGLNVP-HYGWQGLWTWGWAKTQDPEPASSATI 630
Db    446 ILSASKSFSAITATTSSSTVTPPETNLKNYTPPTHYGYQGNWTVTW------KQGSSAQ- 498
```

```
Qy    631 TDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPS--- 687
                  :|  |||   || |:|:    |: |||||     | :::|  ::: :
Db    499 ---------EKTATLTWEQTGYSPNPERVGSLVPNTLWG-AFSDTRAIQNLMDISVNGAD 548

Qy    688 -DHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQT---HTFSLKFSQTYTK 743
           ||      |  : :   :     |   |:||: |: | ||       |  | | : |
Db    549 YSRGFW---VSSLANFLNKSGSDTKRKFRHHSAGYALGVYA-QTPSDDVCSAAFCQLFGK 604

Qy    744 LNERYAKNNVSSKNYSCQGEMLFS-----------LQ-----EGFLLTKLVGLYSYGDHN 787
          : :   | ||  |:  | : :              ||     |  |:       | :  :|
Db    605 DKDYFVSKN-SSTIYA--GSIYYQHISYWNTWNTLLQNTLGAEAPLVLNAQLTYCHASNN 661

Qy    788 CH----HFYTQGENLTS---QGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYS 840
              : ||  :|:|    :| : :   |       |::       : :||:    :::
Db    662 MKTNMTNTYTP-KNVTPSEIKGDWGNDCFGVEFGAKAPIETASLLFDMYSPFVKLQLVHA 720

Qy    841 SLSHFTEVGAYPRSFSTKTPLINVLVPIGVK-GSFMNATQRPQAWTVELAYQPVLYRQEP 899
            | |  :    :      | |: :|||||   | :  :   :: : ||| | : |  |
Db    721 HQDDFKENNSDQGRYFESNNLTNLSMPIGVKLEKFSH--KDTASYNLTLAYAPDIVRSNP 778

Qy    900 GIATQLLAS--KGIWFGSGSPSSRHA 923
              || |    :|    :  :|||
Db    779 DCTASLLVSPTSAVWVTKANNLARHA 804


RESULT 6
Q9Z881
ID   Q9Z881      PRELIMINARY;      PRT;   392 AA.
AC   Q9Z881;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Polymorphic outer membrane protein.
GN   PMP_17.3 OR CPN0470.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CWL029;
RX   MEDLINE=99206606; PubMed=10192388;
RA   Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA   Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT   "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL   Nat. Genet. 21:385-389(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
DR   EMBL; AE001631; AAD18611.1; -.
DR   EMBL; AP002546; BAA98676.1; -.
DR   InterPro; IPR003357; OMP.
DR   Pfam; PF02385; OMP; 1.
KW   Complete proteome.
SQ   SEQUENCE   392 AA;  44365 MW;  53F20CCA2000AE49 CRC64;

  Query Match             10.1%;  Score 512;  DB 16;  Length 392;
  Best Local Similarity   32.1%;  Pred. No. 3e-22;
  Matches  128;  Conservative   68;  Mismatches  165;  Indels   38;  Gaps   10;

Qy    579 PMLSISEASDNQLRSDDMDFSGLNV-PHYGWQGLWTWGWAKTQDPEPASSATITDPQKAN 637
          |:| : : :  :: : ::    :|:  |||:||:|:  | :|      :: : | |:: |
Db     10 PLLYLLDVTAKKIDTSNLIVEAMNLDEHYGYQGIWSPYWMET-----TTTTSSTVPEQTN 64

Qy    638 RFHRTLLLTWLPAGYVPSPKHRSPLIANTLW--------GNMLLATESLKN---SAELTP 686
            || | : | | || |:|:     ||||||        |  :|  ::||     | |
Db     65 TNHRQLYVDWTPVGYRPNPERHGEFIANTLWQSAYNALLGIRILPPQNLKEHDLEASLQ- 123

Qy    687 SDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIA--GQTHTFSLKFSQTYTKL 744
                    |||::: |  ||   ||   ::||:|   |      |:||| |:| ::|
Db    124 -----------GLGLLINQHNREGRKGFRNHTTGYAATTSAKTAARHSFSLGFAQMFSKT 172

Qy    745 NERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGLYSYGDHN--CHHFYTQGENLTSQG 802
           || : :  || ||          |   |: | |   ||||||:  ||  ||:    :|:
Db    173 RERQSPSTTSSHNYFAGLRFDSLLFRDFISTGLSLGYSYGDHHMLCH--YTEILKGSSKA 230

Qy    803 TFRSQTMGGAVFFDLPMKPFGSTHIL-TAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPL 861
           | : |:  ::  |     |   |  |   ||: |: :  | : | | | : | |  | ||
Db    231 FFNNHTLVASL--DCTFLPARITRTLELQPFISAIALRCSQASFQETGDHIRKFHPKHPL 288

Qy    862 INVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSSR 921
           ::  ||| :  :  :   |  || |::| | |||: | : | || | | |    :| |
Db    289 TDLSSPIGFRSEWKTSHHIPMLWTTEISYVPTLYRKNPEMFTTLLISNGTWTTQATPVSY 348

Qy    922 HAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGE 960
          :::: ||   :|  | :||   |    ||||   ||  |
Db    349 NSVAAKIKNTSQLFSRVTLSLDYSAQVSSSTVGQYLKAE 387


RESULT 7
Q8VU50
ID   Q8VU50      PRELIMINARY;      PRT;   700 AA.
AC   Q8VU50;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Putative polymorphic membrane protein (Fragment).
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=POS;
RA   Laroucau K., Souriau A., Rodolakis A.;
RT   "Isolation of a new pmp sequence and evidence of pmp polymorphism in
RT   serotype-1 Chlamydia psittaci strains.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF243415; AAL36959.1; -.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   InterPro; IPR003357; OMP.
DR   Pfam; PF02415; DUF145; 1.
DR   Pfam; PF02385; OMP; 1.
FT   NON_TER     700    700
SQ   SEQUENCE   700 AA;  74200 MW;  2C38E504AF48D1D1 CRC64;

  Query Match             9.4%;  Score 477;  DB 2;  Length 700;
  Best Local Similarity   25.3%;  Pred. No. 7.5e-20;
  Matches  201;  Conservative  109;  Mismatches  264;  Indels  220;  Gaps   40;

Qy      1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSV--PDPTKESLSNKISLTGDTHNLTNCYL 58
            ||   ::||| :||            |  ||:     | |:|::  |  |   |:|
Db      1 MKHPVYWFLISSSL-----------LASNSLSFAQVTNETLTSSDSYNG---NVT---- 41

Qy     59 DNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTV 118
                                 :|     :|    ||      | |    | ||  :||
Db     42 ----------------------SDEFEVKETTSGAIY-----TCEGNVCISYAGKDSPL- 73

Qy    119 EIRDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSY 178
                    | :|    |                                     :| |:
Db     74 ------------NKSCFSETT---------------------------------ENLSF 87

Qy    179 VRGGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAI-YAGTSNSFESNNCDLFFIN 237
          :  |               ||   |||   | ||   |||  :|   :   :   ||
Db     88 IGNG------------YTLCF---DNI---TTTASNPGAINVSGDQKTLNVSGFSLFSCA 129

Qy    238 NACCAG----GAIFSP-ICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTT---- 288
          : |  |    |||  :  : : :||  | : ::| |            :|||||  |
Db    130 H-CPPGTTGYGAIQTKGVSTFSGN--NKLIFDNNC---------STGEGGAIKCATGSNA 177

Qy    289 RLDVTGNRGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINN---IANNKGGAIYIDGT 345
```

```
Db    178 ELKLEGN-SYVVFSGNSSQKKGGAIYTKKLTITADGPTLFSNNSVSASSPKGGAICLDDT 236

Qy    346 SNS-KISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNL 404

Db    237 SSECSLTANLGDITFDGN---KVIKTNGGSSTV---KRNAIDLGSGGKFTKLNAKEGFGI 290

Qy    405 IFYDPIEVSNAGVSVSFNK-EAD--QTGSVVFSGATVNSAD-FHQRNLQTKTPAPLTLSN 460

Db    291 FFHDPIANTGGSTEIELNKTESDTTYTGKIVFSGEKLSDEEKTVPDNLKSYFKQPLKIGA 350

Qy    461 GFLCIEDHAQLTVNRFTQT-GGVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSI 519

Db    351 GSLVLKDGVTLEAKKITQTKGSTVVMDLGTTLQT------PSSSGETITLTNLDINIASL 404

Qy    520 LKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQP 579

Db    405 GGGGGTAPA-----------KLATNTASQAISIAAVNLVNTDSNT-YE----------DP 442

Qy    580 MLSISEA--------SDNQLRSDDMDFSGLNVP-HYGWQGLWTWGWAKTQDPEPASSATI 630

Db    443 ILSASKSFSAITATTSSSTVTPPETNLKNYTPPTHYGYQGNWTLAW------PPGETMQL 496

Qy    631 TDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNM--LLATESLK----NSAEL 684

Db    497 ----------KTATLNWEQTGYSPNPERVGSLVPNTLWGSFSDIRAIQNLMDVSVNGADY 546

Qy    685 TPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHT---FSLKFSQTY 741

Db    547 S---RGFW---VSGLANFLNKSGSDTKRKFRHNSAGYALGVYA-QTPSEDIFSAAFCQLF 599

Qy    742 TKLNERYAKNNVSS 755

Db    600 GKDKDYFLSKNSST 613


RESULT 8
Q9RB61
ID   Q9RB61      PRELIMINARY;      PRT;   347 AA.
AC   Q9RB61;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Polymorphic outer membrane protein.
GN   PMP_17.2.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
DR   EMBL; AP002546; BAA98675.1; -.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   Pfam; PF02415; DUF145; 1.
SQ   SEQUENCE   347 AA;  37314 MW;  8A62367A2664972A CRC64;

  Query Match             8.3%;  Score 422;  DB 16;  Length 347;
  Best Local Similarity   32.4%;  Pred. No. 4.7e-17;
  Matches  110;  Conservative   56;  Mismatches  151;  Indels   22;  Gaps    7;

Qy    246 IFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNI 305

Db      1 MFSALGVIISSNKEIIEISNHSASSINTASGKLYPGGGGIMCTSLVIENNPKGLIFNNKT 60

Qy    306 TKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDG----TSNSKISADRHAIIFNE 361

Db     61 AALSGGAIHTRSFIFQNNGPTAFINNSATSGGALINLSGIGSTPQNFFLSADYGDILFNN 120

Qy    362 NIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVS-VS 420

Db    121 NTI--------TSSSPQPGYRNAL-YAAPGINLKLGARQGYKILFYDPIDHDQTTTDPIV 171

Qy    421 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 480

Db    172 FNYEPHHLGTVLFSGINVDSNATNPLNFLSKFSNSSRLERGVLAIEDRAAISCKTLSQTG 231

Qy    481 GVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNY 540

Db    232 GILRLGNAALIR--TKGPG-----SSINFNAIAINLPSILQSEASAPKFWIYPTLTGSTY 284

Qy    541 TADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQP 579

Db    285 SEDTSSTITLSG-PLTFLNDENENPYDSLDLSEPRKDIP 322


RESULT 9
Q93QE8
ID   Q93QE8      PRELIMINARY;      PRT;  1530 AA.
AC   Q93QE8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Cytolysin.
GN   CTC.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=434/BU;
RA   Lampe M.F., Ballweber L.M., Johnson M.L., Kiselev A.O., Yates J.R.,
RA   Stamm W.E.;
RT   "Cloning and analysis of ctc, a putative cytolysin and member of a
RT   multigene family from Chlamydia trachomatis serovar L2.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF268092; AAK69391.1; -.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   InterPro; IPR001899; Gram_pos_anchor.
DR   InterPro; IPR003357; OMP.
DR   InterPro; IPR000531; TonB_boxC.
DR   Pfam; PF02415; DUF145; 1.
DR   Pfam; PF02385; OMP; 1.
DR   PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
DR   PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
SQ   SEQUENCE   1530 AA;  160459 MW;  F9FCE3F8EF89A89F CRC64;

  Query Match             7.6%;  Score 387;  DB 2;  Length 1530;
  Best Local Similarity   20.5%;  Pred. No. 4e-14;
  Matches  287;  Conservative  140;  Mismatches  422;  Indels  552;  Gaps   63;

Qy      7 FFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTN----------C 56

Db    240 FFVTG-SLSG------EKSLYMP--AGDMVVANCDGAISFEGNSANFANGGAIAASGKVL 290

Qy     57 YLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQK------------------------ 91

Db    291 FVANDKK-TSFIENRALSGGAIAASSDIAFQNCAELVFKGNCAIGTEDKGSLGGGAISSL 349

Qy     92 --------EGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNP 143

Db    350 GTVLLQGNHGITCDKNESASQGGAI-----FGKNCQISDNEGPVVFRDSTAC-------- 396

Qy    144 NAAVNKIREGGAIHAQNLY-INHNHDVVGF-----------MKNFSYVRG--------- 181

Db    397 -------LGGGAIAAQEIVSIQNNQAGISFEGGKASFGGGIACGSFSSAGGASVLGTIDI 449

Qy    182 ----GAISTANTFV----------------------VSENQSCFLFMDNICIQTNTAGK- 214

Db    450 SKNLGAISFSRTLCTTSDLGQMEYQGGGALFGENISLSENAGVLTFKDNIVKTFASNGKI 509
```

```
Qy    215 --GGAIYAGTSNSFESNNCDLFFINNA-------------------------------- 239
          ||||  |       :|:  : |  ||
Db    510 LGGGAILATGKVEITNNSGGISFTGNARAPQALPTQEEFPLFSKKEGRPLSSGYSGGGAI 569

Qy    240 ------------------------------CCAGGAIFS-PICSLTGNRGNIVFYNNRC 267
                                        || |||:      |: ||  :: | ||
Db    570 LGREVAILHNAAVVFEQNRLQCSEEEATLLGCCGGGAVHGMDSTSIVGN-SSVRFGNNYA 628

Qy    268 F-----------KNVETASSEASD---------GGAIKVT-------------------- 287
                  | |: | : : |          |||:: :
Db    629 MGQGVSGGALLSKTVQLAGNGSVDFSGNIASLGGGALQASEGNCELVDNGYVLFRDNRGR 688

Qy    288 --------TRLDV--TGNRGRIFFSDNI-TKNY--------------------------- 309
                  | ||  :||:||: | ||| |: |
Db    689 VYGGAISCLRGDVVISGNKGRVEFKDNIATRLYVEETVEKVEEVEPAPEQKDNNELSFLG 748

Qy    310 ---------------------------------------GGAIYAPVVTLVDN-GPTYF 328
                                                  ||||:|  | :|||     |
Db    749 SVEQSFITAANQALFASEDGDLSPESSISSEELAKRRECAGGAIFAKRVRIVDNQEAVVF 808

Qy    329 INNIANNKGGAIYIDG-TSNSKISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITV 387
           || ::  ||||:       |:      :: : |    | : | :    : |      :
Db    809 SNNFSDIYGGAIFTGSLREEDKLDGQIPEVLISGNAGDVVFSGNSSKRDEHLPHTGGGAI 868

Qy    388 ASSSGEILLGAGSSQNLIFYDPIEVSNAGVSV--------------------SFNKEAD 426
           : :  |    |   |::||: : |   | :                      ||   |
Db    869 CTQNLTISQNTG---NVLFYNNVACSGGAVRIEDHGNVLLEAFGGDIVFKGNSSF--RAQ 923

Qy    427 QTGSVVFSG-----ATVNSAD-----FHQ----RNLQTKTPAPLTLSNG----------- 461
           : :: |:|       :|: :     ||      ||: :  | : | |
Db    924 GSDAIYFAGKESHITALNATEGHAIVFHDALVFENLKERKSAEVLLINSRENPGYTGSIR 983

Qy    462 FL--------CIEDHAQLTVNRFTQTGGVVSLGNGAVLSCY--KNGAGNS---ASNASIT 508
           ||         ||  | |          | : | ||| |  |  |  ||     |: : :
Db    984 FLEAESKVPQCI--HVQ---------QGSLELLNGATLCSYGFKQDAGAKLVLAAGSKLK 1032

Qy    509 LKHIGLNLS--SILKSGAEIP---------LLWVE-------PTNNSNNYTADTAATFSL 550
           :  |  :  :| |  |||           ||:        |  : :  : |  |:||
Db   1033 ILDSGTPVQGHAISKPEAEIESSSEPEGAHSLWIAKNAQTTVPMVHIHTISVD-LASFSS 1091

Qy    551 S---------------------DVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASD 588
           |                     :: | |::  |   ||:  |    :  |::|   :||
Db   1092 SQQEGTVEAPQDIVPGGSYVRSGELNLELVNTTGTG-YENHALLKNEAKVPLMSFVASSD 1150

Qy    589 NQLRSDDMDFSGLNVPH--------------YGWQGLWTWGWAKTQDPEPASSATITDPQ 634
                : | |:|                 ||  |   |  || ||
Db   1151 EA----SAEISNLSVSDLQIHVATPEIEEDTYGHMG--DWSEAKIQD------------- 1191

Qy    635 KANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLW--GNMLLATESLKNSAELTPSDHPF- 691
           ||:: | | ||   |:    | | || | :| | :: :  ||     |
Db   1192 ------GTLVINWNPTGYRLDPQKAGALVFNALWEEGAVLSALKNARFAHNLTAQRMEFD 1245

Qy    692 -----WGITGGGLGMMVYQDPREN---HPGFHMRSSGYSAGMIAGQTHTFSLKFSQT--Y 741
                ||   ||   :      ||      |:     | |||:       | |  |
Db   1246 YSTNVWGFAFGGFRTL----SAENLVAIDGYKGAYGGASAGVDIQLMEDFVLGVSGAAFL 1301

Qy    742 TKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTK--LVGLYSYGDHNCHHFYTQGENLT 799
           |:: :     || |       :: |:  |||       | || |:        || :  |
Db   1302 GKMDSQKFDAEVSRKG------VVGSVYTGFLAGSWFFKGQYSLGE-------TQNDMKT 1348

Qy    800 SQGTF----RSQTMGGAV----------------FFDLPMKPFGSTHILTAPFLGALGI 838
            |       | |  | :                        |: |   |:      :  | |
Db   1349 RYGVLGESSASWTSRGVLADALVEYRSLVGPVRPTFYALHFNPYVEVSYASMKFPG---- 1404

Qy    839 YSSLSHFTEVGAYPRSFSTKTPLINVLVPIG-------VKGSF--MNATQRPQAWTVELA 889
               ||| |   |||   : | |: :|:|       :|| |  :|:      ||
Db   1405 ------FTEQGREARSFEDAS-LTNITIPLGMKFELAFIKGQFSEVNSLGISYAWEA--- 1454

Qy    890 YQPVLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSW----LTLHFQYH 945
               ||:  | | ||| :   | |:     |  :   :   |:  |:     | |
Db   1455 -----YRKVEGGAVQLLEAGFDWEGAPMDLPRQELRVALENNTEWSSYFSTVLGLTAFCG 1509

Qy    946 GFYSSSTFCNY-LNGEIALRF 965
```

```
          || |: :   |  |  : | |
Db   1510 GFTSTDSKLGYEANAGLRLIF 1530


RESULT 10
Q8VU49
ID   Q8VU49      PRELIMINARY;      PRT;   602 AA.
AC   Q8VU49;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Putative polymorphic membrane protein (Fragment).
OS   Chlamydia psittaci (Chlamydophila psittaci).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83554;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LLG;
RA   Laroucau K., Souriau A., Rodolakis A.;
RT   "Isolation of a new pmp sequence and evidence of pmp polymorphism in
RT   serotype-1 Chlamydia psittaci strains.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF243418; AAL36962.1; -.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   InterPro; IPR003357; OMP.
DR   Pfam; PF02415; DUF145; 1.
DR   Pfam; PF02385; OMP; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   602 AA;  65562 MW;  CA486CFACEC131E2 CRC64;

  Query Match             6.9%;  Score 349;  DB 2;  Length 602;
  Best Local Similarity   25.3%;  Pred. No. 1.9e-12;
  Matches  167;  Conservative   97;  Mismatches  261;  Indels  136;  Gaps   36;

Qy    357 IIFNEN-IVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPI-EVSN 414
          | |: | |:|       || |::  :||:|:: |      | |     : |||||    :
Db      4 ITFDGNKIIT-------TSRSSSTVKRNSISLGSGGKFTKLNAKEGFGIFFYDPIANTGD 56

Qy    415 AGVSVSFNK----EADQTGSVVFSGATVNS-----ADFHQRNLQTKTPAPLTLSNGFLCI 465
              :  ||         || :|||| ::      ||    ||::    || :  | | :
Db     57 TNTEIELNKAEGGSTTYTGKIVFSGEKLSDEEKKVAD----NLKSYFTQPLKIGAGSLVL 112

Qy    466 EDHAQLTVNRFTQT-GGVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGA 524
          :|   |   : :|| |  | :   |  |          |:|  :||| :: :|::|:   |
Db    113 KDGVTLEAKKVSQTDGSTVVMDLGTTLQ-------TSSSGETITLTNLDINVASLGGGGV 165

Qy    525 EIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLS-I 583
                ||   :    | :            ::|:|  ||: ||   |    :|||  : |
Db    166 APDPAKVEAQASGKTVTINA----------VNLVDTDGNA-YEYPILA---TSQPFTAII 211

Qy    584 SEASDNQLRSDDMDFSGLNVP--HYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHR 641
          ::|  :   :   |       |  |||:|| ||  |      :  :||
Db    212 AKAGSSGTTTPTPDNLKNYTPPTHYGYQGNWTVTW------KLGTSAQ----------EE 255

Qy    642 TLLLTWLPAGYVPSPKHRSPLIANTLWGNM--LLATESL----KNSAELTPSDHPFWGIT 695
          |  |||    | |:|: : ||: |||||:   : | ::|     | |:       ||
Db    256 TATLTWEQTDYSPNPERQGPLVPNTLWGSFSDIRAIQNLIDISVNGADYR---RGFW--- 309

Qy    696 GGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHT---FSLKFSQTYTKLNERYAKNN 752
            |||  :::        |   |:||: |: | || |    ||  | | : |  : :   |
Db    310 VSGLGNFLHKSGSNTKRKFRHHSAGYALGVYA-QTSTEDVFSAAFCQLFGKDKDYFVSKN 368

Qy    753 VSSKNYSCQGEMLF---SLQEGF--LLTKLVGL---------YSYGDHNCHHFYTQGENL 798
           ||  |:  | : :   |    :  ||   :|           :|    ||       |:
Db    369 -SSNIYA--GSIYYQHISYWNAWQNLLQSTIGAEAPLVLNAQLTY----CHASNNMKTNM 421

Qy    799 TS-----------QGTFRSQTMG---GAVFFDLPMK-PFGSTHILTAPFLGALGIYSSL 842
          |:           :| : :   |   ||:    |:: |        :|||    :::
Db    422 TNTYVPKNVTLSEIKGDWGNDCFGVEFGAM---APIETPSSFLFDRYSPFLQLQLVHAHQ 478

Qy    843 SHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIA 902
            | |  :    :  | |: :|||:|     |    :: :  || | :  |   |
Db    479 DDFKENNSDQGRYFESSNLTNLSMPIGIKFERF-AYNDVASYHLTAAYAPDIVRSNPDCT 537
```

```
Qy    903 TQLLAS--KGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLH-----FQYHGFY---SSST 952
          || |    :|    :  :| |.        |  ::| |      |   ||    || |
Db    538 ASLLVSPTSAVWVTKANNLARSAF------MLQAGNYLALSHNMELFSQFGFEIRGSSRT 591

Qy    953 F 953
          :
Db    592 Y 592
```

RESULT 11
Q8VU48
ID Q8VU48 PRELIMINARY; PRT; 581 AA.
AC Q8VU48;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative polymorphic membrane protein (Fragment).
OS Chlamydia psittaci (Chlamydophila psittaci).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83554;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POS;
RA Laroucau K., Souriau A., Rodolakis A.;
RT "Isolation of a new pmp sequence and evidence of pmp polymorphism in
RT serotype-1 Chlamydia psittaci strains.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF243419; AAL36963.1; -.
DR InterPro; IPR003368; Chlamydia_PMP.
DR InterPro; IPR003357; OMP.
DR Pfam; PF02415; DUF145; 1.
DR Pfam; PF02385; OMP; 1.
FT NON_TER 1 1
SQ SEQUENCE 581 AA; 62860 MW; CDDF3C98522E112F CRC64;

Query Match 6.8%; Score 346.5; DB 2; Length 581;
Best Local Similarity 24.9%; Pred. No. 2.5e-12;
Matches 155; Conservative 95; Mismatches 251; Indels 121; Gaps 31;

```
Qy    357 IIFNEN-IVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPI-EVSN 414
          | |: | |:|       || |::  :||:|:: |      | |     : |||||    :
Db      4 ITFDGNKIIT-------TSRSSSTVKRNSISLGSGGKFTKLNAKEGFGIFFYDPIANTGD 56

Qy    415 AGVSVSFNK----EADQTGSVVFSGATVNS-----ADFHQRNLQTKTPAPLTLSNGFLCI 465
            :  ||        || :|||| ::      ||    ||::    || :  | | :
Db     57 TNTEIELNKAEGGSTTYTGKIVFSGEKLSDEEKKVAD----NLKSYFTQPLKIGAGSLVL 112

Qy    466 EDHAQLTVNRFTQTGG---VVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSILKS 522
          :|   |   : :|| |   |: ||              |:|  :||| :: :|::|:
Db    113 KDGVTLEAKKVSQTDGSTAVMDLG---------TTLQTSSSGETITLTNLDINVASLGGG 163

Qy    523 GAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLS 582
          |       ||   :    | :           ::|:|  ||: ||   |    :|||  :
Db    164 GVAPDPAKVEAQASGKTVTINA----------VNLVDTDGNA-YEYPILA---TSQPFTA 209

Qy    583 -ISEASDNQLRSDDMDFSGLNVP--HYGWQGLWTWGWAKTQDPEPASSATITDPQKANRF 639
           |::|  :   :   |      |  |||:|| ||  |      :  :||
Db    210 IIAKAGSSGTTTPTDNLKNYTPPTHYGYQGNWTVTW------KLGTSAQ---------- 253

Qy    640 HRTLLLTWLPAGYVPSPKHRSPLIANTLWGNM--LLATESLK----NSAELTPSDHPFWG 693
          |  | |   || |:|:    |: |||||:   : | ::|     | |: :     ||
Db    254 EETATLNWEQTGYSPNPERVGSLVPNTLWGSFSDIRAIQNLMDVSVNGADYS---RGFW- 309

Qy    694 ITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHT---FSLKFSQTYTKLNERYAK 750
            ||  : :   :   |  |:||: |: | || :   ||  | | : |  : :
Db    310 --VSGLANFLNKSGSDTKRKFRHNSAGYALGVYA-QTPSEDIFSAAFCQLFGKDKDYFLS 366

Qy    751 NNVSSKNYSCQGEMLF---SLQEGF--LLTKLVGL---------YSYGDHNCHHFYTQGE 796
          | ||  |:  | : :   |    :  ||   :|           :|    ||
Db    367 KN-SSTIYA--GSIYYQHISYWNAWQNLLQNTIGAEAPLVLNAQLTY----CHASNNMKT 419

Qy    797 NLTS------------QGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSH 844
```

```
          |:|:            :| : |   |       |::       : :||:    :::
Db    420 NMTNTYTPKNVTPSEIKGDWGSDCFGVEFGAKAPIETASLLFDMYSPFVKLQLVHAHQDD 479

Qy    845 FTEVGAYPRSFSTKTPLINVLVPIGVK-GSFMNATQRPQAWTVELAYQPVLYRQEPGIAT 903
          | |  :    :     | |: :|||||   | :  :   :: : ||| | : |  |
Db    480 FKENNSDQGRYFESNNLTNLSMPIGVKLEKFSH--EDTASYNLTLAYAPDIVRSNPDCTA 537

Qy    904 QLLAS--KGIWFGSGSPSSRHA 923
          || |    :|    :  :|||
Db    538 SLLVSPTSAVWVTKANNLARHA 559
```

RESULT 12
Q8VL57
ID Q8VL57 PRELIMINARY; PRT; 601 AA.
AC Q8VL57;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative polymorphic membrane protein (Fragment).
OS Chlamydia psittaci (Chlamydophila psittaci).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83554;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=POS, AND LLG;
RA Laroucau K., Souriau A., Rodolakis A.;
RT "Isolation of a new pmp sequence and evidence of pmp polymorphism in
RT serotype-1 Chlamydia psittaci strains.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF243416; AAL36960.1; -.
DR EMBL; AF243417; AAL36961.1; -.
DR InterPro; IPR003368; Chlamydia_PMP.
DR InterPro; IPR003357; OMP.
DR Pfam; PF02415; DUF145; 1.
DR Pfam; PF02385; OMP; 1.
FT NON_TER 1 1
SQ SEQUENCE 601 AA; 65476 MW; D6AA97EC9072C757 CRC64;

Query Match 6.8%; Score 344.5; DB 2; Length 601;
Best Local Similarity 25.0%; Pred. No. 3.4e-12;
Matches 166; Conservative 96; Mismatches 260; Indels 141; Gaps 36;

```
Qy    357 IIFNEN-IVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPI-EVSN 414
          | |: | |:|       || |::  :||:|:: |      | |     : |||||    :
Db      4 ITFDGNKIIT-------TSRSSSTVKRNSISLGSGGKFTKLNAKEGFGIFFYDPIANTGD 56

Qy    415 AGVSVSFNK----EADQTGSVVFSGATVNS-----ADFHQRNLQTKTPAPLTLSNGFLCI 465
            :  ||        || :|||| ::      ||    ||::    || :  | | :
Db     57 TNTEIELNKAEGGSTTYTGKIVFSGEKLSDEEKKVAD----NLKSYFTQPLKIGAGSLVL 112

Qy    466 EDHAQLTVNRFTQT-GGVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGA 524
          :|   |   : :|| |  | : |  |          |:|  :||| :: :|::|:   |
Db    113 KDGVTLEAKKVSQTDGSTVVMDLGTTLQ-------TSSSGETITLTNLDINVASLGGGGV 165

Qy    525 EIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLS-I 583
            ||   :    | :           ::|:|  ||: ||   |    :|:|  : |
Db    166 APDPAKVEAQASGKTVTINA----------VNLVDTDGNA-YEYPILA---TSKPFTAII 211

Qy    584 SEASDNQLRSDDMDFSGLNVP--HYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHR 641
          ::|  :   :   |      |  |||:|| ||  |      :  :||
Db    212 AKAGSSGTTTPTDNLKNYTPPTHYGYQGNWTVTW------KLGTSAQ----------EE 255

Qy    642 TLLLTWLPAGYVPSPKHRSPLIANTLWGNM--LLATESL----KNSAELTPSDHPFWGIT 695
          |  |||    | |:|: : ||:  ||||:   : | ::|     | |:       ||
Db    256 TATLTWEQTDYSPNPERQGPLVPKTLWGSFSDIRAIQNLIDISVNGADYR---RGFW--- 309

Qy    696 GGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHT---FSLKFSQTYTKLNERYAKNN 752
            |||  :::        |   |:||: |: | || |   ||  | | : |  : :   |
Db    310 VSGLGNFLHKSGSNTKRKFRHHSAGYALGVYA-QTSTEDVFSAAFCQLFGKDKDYFVSKN 368

Qy    753 VSSKNYSCQGEMLF---SLQEGF--LLTKLVGL---------YSYGDHNCHHFYTQGENL 798
          ||  |:  | : :   |    :  ||   :|           :|    ||       |:
```

```
Db    369 -SSNIYA--GSIYYQHISYWNAWQNLLQSTIGAEAPLVLNAQLTY----CHASNNMKTNM 421

Qy    799 TS------------QGTFRSQTMG---GAVFFDLPMKPFGSTHIL---TAPFLGALGIYS 840
Db    422 TNTYVPKNVTLSEIKGDWGNDCFGVEFGA------MAPIENASFLFDRYSPFLQLQLVHA 475

Qy    841 SLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPG 900
Db    476 HQDDFKENNSDQGRYFESSNLTNLSMPIGIKFERF-AYNDVASYHLTAAYAPDIVRSNPD 534

Qy    901 IATQLLAS--KGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLH-----FQYHGFY---SS 950
Db    535 CTASLLVSPTSAVWVTKANNLARSAF------MLQAGNYLALSHNMELFSQFGFEIRGSS 588

Qy    951 STF 953
Db    589 RTY 591


RESULT 13
Q9RB68
ID   Q9RB68      PRELIMINARY;      PRT;   494 AA.
AC   Q9RB68;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Polymorphic outer membrane protein G family.
GN   PMP_5_1.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
DR   EMBL; AP002545; BAA98230.1; -.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   Pfam; PF02415; DUF145; 2.
SQ   SEQUENCE   494 AA;  50229 MW;  5BD8A44AB3A64171 CRC64;

  Query Match             6.3%;  Score 321;  DB 16;  Length 494;
  Best Local Similarity   26.0%;  Pred. No. 6.2e-11;
  Matches  128;  Conservative   75;  Mismatches  186;  Indels  104;  Gaps   20;

Qy     33 PDPTKESLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKE 92
Db     45 PKSTTDAAGTTYSLTGEV-----LYID------------PGKGGSITGT---CFVETAGD 84

Qy     93 GIYFA-----KNLTPESGGAIGYAS-PNSPTVEIRDTIGPVIFENNTCCRPFTSSNPNAA 146
Db     85 LTFLGNGNTLKFLSVDAGANIAVAHVQGSKNLSFTDFLSLVITE-----------SPKSA 133

Qy    147 VN-------------------------KIREGGAIHAQNLYINHNHDVVGFMKNFSYVR 180
Db    134 VTTGKGSLVSLGAVQLQDINTLVLTSNASVEDGGVIKGNSCLIQGIKNSAIFGQNTSSKK 193

Qy    181 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNA- 239
Db    194 GGAISTTQGLTIENNLGTLKFNENKAVTS-----GGALDLGAASTFTANH-ELIFSQNKT 247

Qy    240 ---CCAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNR 296
Db    248 SGNAANGGAI---NCS-----GDLTFTDNTSLLLQE--NSTMQDGGALCSTGTISITGS- 296

Qy    297 GRIFFSDNITKNYGGAIYAPVVTLV-DNGPTYFINNI---ANNKGGAIYIDGTSNSKISA 352
Db    297 DSINVIGNTSGQKGGAISAASLKILGGQGGALFSNNVVTHATPLGGAIFINTGGSLQLFT 356

Qy    353 DRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEV 412
Db    357 QGGDIVFEGNQVT-TTAPNATT------KRNVIHLESTAKWTGLAASQGNAIYFYDPITT 409

Qy    413 SNAGVS-------VSFNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCI 465
Db    410 NDTGASDNLRINEVSANQKL--SGSIVFSGERLSTAEAIAENLTSRINQPVTLVEGSLVL 467

Qy    466 EDHAQLTVNRFTQ 478
Db    468 KQGVTLITQGFSQ 480


RESULT 14
Q9RB70
ID   Q9RB70      PRELIMINARY;      PRT;   427 AA.
AC   Q9RB70;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Polymorphic outer membrane protein G family.
GN   PMP_4_1.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
DR   EMBL; AP002545; BAA98227.1; -.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   Pfam; PF02415; DUF145; 1.
SQ   SEQUENCE   427 AA;  43419 MW;  AB4BBBC1594DD2B1 CRC64;

  Query Match             5.8%;  Score 297;  DB 16;  Length 427;
  Best Local Similarity   30.0%;  Pred. No. 1.3e-09;
  Matches  119;  Conservative   47;  Mismatches  121;  Indels  110;  Gaps   18;

Qy     61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120
Db    130 LRMLAA--PRTTGKG-AIKITDGLVF--ESIGNLDLNENASSENGGAI-----NTKTLSL 179

Qy    121 RDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQ-NLYINHNHDVVGFMKNFSYV 179
Db    180 TGSTRFVAFLGN--------------SSSQQGGAIYASGDSVISENAGILSFGNNSATT 224

Qy    180 RGGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNA 239
Db    225 SGGAISAEGNLVISNNQN--IFFDGCKATTN----GGAI-------------------D 258

Qy    240 CCAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRI 299
Db    259 CNKAGANPDPILTLSGNE-SLHFLNN----------TAGNSGGAI-YTKKLVLSSGRGGV 306

Qy    300 FFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIF 359
Db    307 LFSNNKAAN-----------------------ATPKGGAIAILDSGEISISADLGNIIF 342

Qy    360 NENIVTNVTNANGTSTSANPP--RRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGV 417
Db    343 ---------EGNTTSTTGSPASVTRNAIDLASNAKFLNLRATRGNKVIFYDPITSSGATD 393

Qy    418 SVSFNKEAD------QTGSVVFSGATVNSADFHQRNL 448
Db    394 KLSLNK-ADAGSGNTYEGYIVFSGEKLSEV----RNL 425


RESULT 15
```

```
Q9RB67
ID   Q9RB67      PRELIMINARY;      PRT;   445 AA.
AC   Q9RB67;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Pmp_5.
GN   PMP_5_2.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
DR   EMBL; AP002545; BAA98231.1; -.
DR   InterPro; IPR003357; OMP.
DR   Pfam; PF02385; OMP; 1.
SQ   SEQUENCE   445 AA;  49353 MW;  EFA24AFC9C5097A6 CRC64;

  Query Match             5.4%;  Score 275.5;  DB 16;  Length 445;
  Best Local Similarity   24.8%;  Pred. No. 2.5e-08;
  Matches  114;  Conservative   68;  Mismatches  171;  Indels  107;  Gaps   22;

Qy     501 SASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDD 560
           ||:| :|||      |: :   ||              |:|   :  :|   ||||
Db      24 SAANKNITLTG---TLALVNADGAFY----------ENHTLQDSQDYSF--VKLS---- 63

Qy     561 YGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSGLNVPHYGWQGLWTWGWAKTQ 620
               | :|   | :|:    | : ::             ||||:|| |
Db      64 ---PGAGGTIITQDASQKPL----EVAPSR-----------PHYGYQGHWN------- 96

Qy     621 DPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNML---LATES 677
               | | | :||       | |:  ||:|:|: :  |: |:|||: :      |
Db      97 --VQVIPGTGTQPSQAN-------LEWVRTGYLPNPERQGSLVPNSLWGSFVDQRAIQEI 147

Qy     678 LKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFS--- 734
           : ||:::   :   |   | |:   :::|    | |:     ||  |:    || ||
Db     148 MVNSSQILCQERGVW---GAGIANFLHRDKINEH-GYRHSGVGYLVGV---GTHAFSDAT 200

Qy     735 --LKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGLY--SYGDHNCHH 790
              | | :::      |: | |||:      :  |::        | |  |  :  |:
Db     201 INAAFCQLFSR-----DKDYVVSKNHGTSYSGVVFLEDTLEFRSPQGFYTDSSSEACCNQ 255

Qy     791 FYTQGENLT----------------SQGTFRSQTMG---GAVFFDLPMKPFGSTHILT- 829
             |    |:                  :||:: :   |   ||  :  |     || :
Db     256 VVTIDMQLSYSHRNNDMKTKYTTYPEAQGSWANDVFGLEFGATTYYYP----NSTFLFDY 311

Qy     830 -APFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVEL 888
            :|||     |:    | | |   | | |   | |: ||||||    :  :|  :: :
Db     312 YSPFLRLQCTYAHQEDFKETGGEVRHF-TSGDLFNLAVPIGVKFERFSDCKR-GSYELTF 369

Qy     889 AYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKI 928
           || | : |::|  :|  |||   |   |:  ||  :  ::
Db     370 AYVPDVIRKDPK-STATLASGATWSTHGNNLSRQGLQLRL 408


Search completed: February  4, 2003, 20:57:49
Job time : 101 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2003, 19:10:17 ; Search time 40 Seconds
(without alignments)
1000.616 Million cell updates/sec

Title: US-09-677-752-2
Perfect score: 5086
Sequence: 1 MKKAFFFFLIGNSLSGLARE.....GFYSSSTFCNYLNGEIALRF 965

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4989.5	98.1	964	1	PMPE_CHLTR	O84877 chlamydia t
2	3648.5	71.7	976	1	PMPE_CHLMU	Q9pl47 chlamydia m
3	1120.5	22.0	938	1	PM15_CHLPN	Q9z883 chlamydia p
4	1060.5	20.9	952	1	PM16_CHLPN	Q9z882 chlamydia p
5	951.5	18.7	946	1	PM18_CHLPN	Q9z880 chlamydia p
6	866.5	17.0	1034	1	PMPF_CHLTR	P38008 chlamydia t
7	793	15.6	1025	1	PMPF_CHLMU	Q9pl46 chlamydia m
8	655.5	12.9	1013	1	PMPG_CHLTR	O84879 chlamydia t
9	644	12.7	936	1	PMP7_CHLPN	Q9z898 chlamydia p
10	620.5	12.2	987	1	PMPG_CHLMU	Q9pl45 chlamydia m
11	617	12.1	930	1	PMP8_CHLPN	Q9z393 chlamydia p
12	597	11.7	928	1	PM11_CHLPN	O86164 chlamydia p
13	595	11.7	1016	1	PMPH_CHLTR	O84880 chlamydia t
14	586.5	11.5	928	1	PM10_CHLPN	Q9rb65 chlamydia p
15	586	11.5	947	1	PM19_CHLPN	Q9z813 chlamydia p
16	585.5	11.5	928	1	PMP9_CHLPN	Q9z398 chlamydia p
17	583	11.5	978	1	PM14_CHLPN	Q9z895 chlamydia p
18	575	11.3	973	1	PM13_CHLPN	Q9z896 chlamydia p
19	570.5	11.2	922	1	PMP1_CHLPN	Q9z9g5 chlamydia p
20	567	11.1	1276	1	PMP6_CHLPN	Q9z899 chlamydia p
21	564	11.1	980	1	PMPH_CHLMU	Q9pl44 chlamydia m
22	546.5	10.7	976	1	PMPA_CHLMU	Q9pjy3 chlamydia m
23	525.5	10.3	841	1	PMP2_CHLPN	Q9z3a1 chlamydia p
24	519.5	10.2	975	1	PMPA_CHLTR	O84417 chlamydia t
25	517	10.2	1723	1	PM20_CHLPN	Q9z812 chlamydia p
26	474.5	9.3	1609	1	PM21_CHLPN	Q9z6u5 chlamydia p
27	470.5	9.3	878	1	PMPI_CHLTR	O84882 chlamydia t
28	457	9.0	867	1	PMPI_CHLMU	Q9pl41 chlamydia m
29	395	7.8	1754	1	PMPB_CHLTR	O84418 chlamydia t
30	389.5	7.7	1672	1	PMPB_CHLMU	Q9pjy2 chlamydia m
31	386	7.6	1531	1	PMPD_CHLTR	O84818 chlamydia t
32	377.5	7.4	1460	1	PMPC_CHLMU	Q9pjy1 chlamydia m
33	358.5	7.0	1520	1	PMPD_CHLMU	Q9plb0 chlamydia m
34	342	6.7	1770	1	PMPC_CHLTR	O84419 chlamydia t
35	249.5	4.9	514	1	PM12_CHLPN	Q9z3d6 chlamydia p
36	220.5	4.3	1250	1	YFAL_ECOLI	P45508 escherichia
37	210.5	4.1	2021	1	OMPA_RICCN	Q52657 rickettsia
38	193	3.8	1643	1	OMPB_RICPR	Q53020 r outer mem
39	191.5	3.8	1645	1	OMPB_RICTY	P96989 r outer mem
40	190	3.7	1300	1	120K_RICRI	P14914 rickettsia
41	190	3.7	1654	1	OMPB_RICRI	Q53047 r outer mem
42	189	3.7	1656	1	OMPB_RICJA	O06653 r outer mem
43	178	3.5	2003	1	YDBA_ECOLI	P33666 escherichia
44	174	3.4	2249	1	OMPA_RICRI	P15921 rickettsia
45	173.5	3.4	1655	1	OMPB_RICCN	Q9kka3 r outer mem

ALIGNMENTS

RESULT 1
PMPE_CHLTR
ID PMPE_CHLTR STANDARD; PRT; 964 AA.
AC O84877;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable outer membrane protein pmpE precursor (Polymorphic membrane
DE protein E).
GN PMPE OR CT869.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UW-3/Cx;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis.";
RL Science 282:754-759(1998).
CC -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; AE001360; AAC68467.1; -.
DR InterPro; IPR003368; Chlamydia_PMP.
DR InterPro; IPR003357; OMP.
DR Pfam; PF02385; OMP; 1.
DR Pfam; PF02415; DUF145; 1.
KW Outer membrane; Signal; Multigene family; Complete proteome.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 964 PROBABLE OUTER MEMBRANE PROTEIN PMPE.
SQ SEQUENCE 964 AA; 104703 MW; 1B998A7D2E571CE2 CRC64;

Query Match 98.1%; Score 4989.5; DB 1; Length 964;
Best Local Similarity 98.3%; Pred. No. 4.7e-306;
Matches 949; Conservative 7; Mismatches 8; Indels 1; Gaps 1;

```
Qy       1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 60

Qy      61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120
```

```
Qy    121 RDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVR 180
          |||||||||||||||| ||  || || :|||||||||||||||||||||||||||||||: 
Db    121 RDTIGPVIFENNTCCRLFTWRNPYAA-DKIREGGAIHAQNLYINHNHDVVGFMKNFSYVQ 179

Qy    181 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    180 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 239

Qy    241 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    240 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 299

Qy    301 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFN 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    300 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFN 359

Qy    361 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    360 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 419

Qy    421 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    420 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 479

Qy    481 GVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNY 540
          |||||||||||||||| |:|||||||||||||||||||||||||||||||||||||||||| 
Db    480 GVVSLGNGAVLSCYKNGTGDSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNY 539

Qy    541 TADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||:|:::|||| 
Db    540 TADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLQSENIDFSG 599

Qy    601 LNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRS 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    600 LNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRS 659

Qy    661 PLIANTLWGNMLLATESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSG 720
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||| 
Db    660 PLIANTLWGNMLLATESLKNSAELTPSGHPFWGITGGGLGMMVYQDPRENHPGFHMRSSG 719

Qy    721 YSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGL 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    720 YSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGL 779

Qy    781 YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYS 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    780 YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYS 839

Qy    841 SLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPG 900
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||| 
Db    840 SLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATHRPQAWTVELAYQPVLYRQEPG 899

Qy    901 IATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGE 960
          || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    900 IAAQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGE 959

Qy    961 IALRF 965
          ||||| 
Db    960 IALRF 964


RESULT 2
PMPE_CHLMU
ID   PMPE_CHLMU     STANDARD;      PRT;   976 AA.
AC   Q9PL47;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable outer membrane protein pmpE precursor (Polymorphic membrane
DE   protein E).
GN   PMPE OR TC0261.
OS   Chlamydia muridarum.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=83560;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MoPn / Nigg;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
CC   -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC       (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE002293; AAF39130.1; -.
DR   TIGR; TC0261; -.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   InterPro; IPR003357; OMP.
DR   Pfam; PF02385; OMP; 1.
DR   Pfam; PF02415; DUF145; 1.
KW   Outer membrane; Signal; Multigene family; Complete proteome.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    976       PROBABLE OUTER MEMBRANE PROTEIN PMPE.
SQ   SEQUENCE   976 AA;  105964 MW;  0BEA56158E3CEEA9 CRC64;

  Query Match             71.7%;  Score 3648.5;  DB 1;  Length 976;
  Best Local Similarity   70.5%;  Pred. No. 9.6e-222;
  Matches  689;  Conservative  121;  Mismatches  154;  Indels   13;  Gaps    3;

Qy      1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 60
          ||| ||| |||:|: |  ||||  | | |   |          |::| |||||||: :|||
Db      1 MKKLFFFVLIGSSILGFTREVPPSILLKPILNPYHMTGLFFPKVNLLGDTHNLTDYHLDN 60

Qy     61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120
          |: ||| ||:|| |||| |:|||| | ||||:||:  |||||||||| ||  : |:||:|
Db     61 LKCILACLQRTPYEGAAFTVTDYLGFSDTQKDGIFCFKNLTPESGGVIGSPTQNTPTIKI 120

Qy    121 RDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVR 180
           :|||||:|||||| | :| ::|    || |||||||| ::||::| ::|||:|||:||:
Db    121 HNTIGPVLFENNTCHRLWTQTDPENEGNKAREGGAIHAGDVYISNNQNLVGFIKNFAYVQ 180

Qy    181 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 240
          ||||| ||||   ||:| || ::| |||| | |||||||  || |||:|| || || |:
Db    181 GGAIS-ANTFAYKENKSSFLCLNNSCIQTKTGGKGGAIYVSTSCSFENNNKDLLFIQNSG 239

Qy    241 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 300
          ||||||||| ||| ||:|:||||:|  ||||: |::|: |||||||||||:| |  :||
Db    240 CAGGAIFSPTCSLIGNQGDIVFYSNHGFKNVDNATNESGDGGAIKVTTRLDITNNGSQIF 299

Qy    301 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFD 360
          |||||::|:||||:|| : || |||||| |||||: |||||| ||  |||||| |||||:
Db    300 FSDNISRNFGGAIHAPCLHLVGNGPTYFTNNIANHTGGAIYITGTETSKISADHHAIIFD 359

Qy    361 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 420
           ||  | |||:|:|:: ||| |||||: :|:| | |||| ||||||||||:|:||||:|
Db    360 NNISANATNADGSSSNTNPPHRNAITMDNSAGGIELGAGKSQNLIFYDPIQVTNAGVTVD 419

Qy    421 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 480
          |||:| ||| |||||||| |||  | ||||||| ||||:| ||||| |||||||||||||
Db    420 FNKDASQTGCVVFSGATVLSADISQANLQTKTPATLTLSHGLLCIEDRAQLTVNNFTQTG 479
```

```
Qy     481 GVVSLGNGAVLSCYKNGAGN--------SASNASITLKHIGLNLSSILKSGAEIPLLWVE 532
           |:|:||||||||| |::   :        : ::||:|| |||||| |||| |||:||||||
Db     480 GIVALGNGAVLSSYQHSTTDATQTPPTTTTTDASVTLNHIGLNLPSILKDGAEMPLLWVE 539

Qy     533 PTN----NSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASD 588
           | :    |:  ||:||||:|||:   |||||: |||||:|||: || :||||:||||||
Db     540 PISTTQGNTTTYTSDTAASFSLNGATLSLIDEDGNSPYENTDLSRALYAQPMLAISEASD 599

Qy     589 NQLRSDDMDFSGLNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWL 648
           |||:|: |||| :|||||||||||||||||::|     ||||||:|||:|||||||||||
Db     600 NQLQSESMDFSKVNVPHYGWQGLWTWGWAKTENPTTTPPATITDPKKANQFHRTLLLTWL 659

Qy     649 PAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSDHPFWGITGGGLGMMVYQDPR 708
           ||||:|||||:|||||||||||:| |||:||||:     | |||||||||||||||:||
Db     660 PAGYIPSPKHKSPLIANTLWGNILFATENLKNSSGQELLDRPFWGITGGGLGMMVYQEPR 719

Qy     709 ENHPGFHMRSSGYSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSL 768
           ::|||||| :|||||||| | ||||||:|||:|||||||||| ||||||||||||| ||
Db     720 KDHPGFHMHTSGYSAGMITGNTHTFSLRFSQSYTKLNERYAKNYVSSKNYSCQGEMLLSL 779

Qy     769 QEGFLLTKLVGLYSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHIL 828
           ||| :||||:||||||:|| |||||||||:|:||| | ||| |||||||||:|||| ||||
Db     780 QEGLMLTKLIGLYSYGNHNSHHFYTQGEDLSSQGEFHSQTFGGAVFFDLPLKPFGRTHIL 839

Qy     829 TAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVEL 888
           |||||||:|:|| || |||||||||:| |:|||||||:|||||||||| |||||||||
Db     840 TAPFLGAIGMYSKLSSFTEVGAYPRTFITETPLINVLIPIGVKGSFMNATHRPQAWTVEL 899

Qy     889 AYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFY 948
           |||||||||| |:||||| ||:||| |||:||||::|||||:|| | : ||  ||||:|
Db     900 AYQPVLYRQEPSISTQLLAGKGMWFGHGSPASRHALAYKISQKTQLLRFATLQLQYHGYY 959

Qy     949 SSSTFCNYLNGEIALRF 965
           |||||||||||::|||
Db     960 SSSTFCNYLNGEVSLRF 976
```

```
RESULT 3
PM15_CHLPN
ID   PM15_CHLPN     STANDARD;      PRT;   938 AA.
AC   Q9Z883;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable outer membrane protein pmp15 precursor (Polymorphic membrane
DE   protein 15).
GN   PMP15 OR CPN0466 OR CP0286.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CWL029;
RX   MEDLINE=99206606; PubMed=10192388;
RA   Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA   Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT   "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL   Nat. Genet. 21:385-389(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AR39;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
CC   -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC       (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001631; AAD18608.1; -.
DR   EMBL; AE002190; AAF38143.1; -.
DR   EMBL; AP002546; BAA98672.1; -.
DR   TIGR; CP0286; -.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   InterPro; IPR003357; OMP.
DR   Pfam; PF02385; OMP; 1.
DR   Pfam; PF02415; DUF145; 1.
KW   Outer membrane; Signal; Multigene family; Complete proteome.
FT   SIGNAL       1     17       POTENTIAL.
FT   CHAIN       18    938       PROBABLE OUTER MEMBRANE PROTEIN PMP15.
SQ   SEQUENCE   938 AA;  102194 MW;  A904AAB05B567455 CRC64;

  Query Match             22.0%;  Score 1120.5;  DB 1;  Length 938;
  Best Local Similarity   30.9%;  Pred. No. 7.7e-63;
  Matches  308;  Conservative  163;  Mismatches  418;  Indels  109;  Gaps   31;

Qy       6 FFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLT---------NC 56
           | |:: |   ||   : : | | |      |           | |||           |
Db      12 FTFVLAN--EGLQLPLETYITLSPEYQAAPQ----------VGFTHNQNQDLAIVGNHND 59

Qy      57 YLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSP 116
           :: : :|         : | |:|  : |         ::| ||: | ||||| ||: |
Db      60 FILDYKYY-------RSNGGALTCKNLL--ISENIGNVFFEKNVCPNSGGAI-YAAQNC- 108

Qy     117 TVEIRDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNF 176
                ||     :|        | || |    :  |||: | |  | :|     |: |
Db     109 ------TISK--NQNYAFTTNLVSDNPTATAGSLL-GGALFAINCSITNNLGQGTFVDNL 159

Qy     177 SYVRGGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFI 236
           :  :|||: |     : :|:   :   |  :  |:   || ||:| | : | |:  :
Db     160 ALNKGGALYTETNLSIKDNKGPIIIKQNRAL--NSDSLGGGIYSGNSLNIEGNSGAIQIT 217

Qy     237 NNACCAGGAIFS-PICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGN 295
           :|:  :|| |||    ::: |:  |    |  | |      |  : ||    ||   :  |
Db     218 SNSSGSGGGIFSTQTLTISSNKKLIEISENSAFAN--NYGSNFNPGGGGLTTTFCTILNN 275

Qy     296 RGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYID-GTSNSK--ISA 352
           |  : |::| ::: ||||:|  : : :||| ||:|| |   |  : :  |: |    :||
Db     276 REGVLFNNNQSQSNGGAIHAKSIIIKENGPVYFLNNTATRGGALLNLSAGSGNGSFILSA 335

Qy     353 DRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEV 412
           |   |||| |          :  : ||| ||||  :: :  : :||     ::||||||
Db     336 DNGDIIFNNN--------TASKHALNPPYRNAIH-STPNMNLQIGARPGYRVLFYDPIEH 386

Qy     413 S-NAGVSVSFNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQL 471
             :  : || |   ||:|:|||  |:     : |  :     |  | | :|| | |
Db     387 ELPSSFPILFNFETGHTGTVLFSGEHVHQNFTDEMNFFSYLRNTSELRQGVLAVEDGAGL 446

Qy     472 TVNRFTQTGGVVSLGNGAVLSCYKNGAGN--------SASNASITLKHIGLNLSSILKSG 523
           :| | || : || |||::     ||          :   ::||| || ::| |||
Db     447 ACYKFFQRGGTLLLGQGAVITT----AGTIPTPSSTPTTVGSTITLNHIAIDLPSILSFQ 502

Qy     524 AEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSI 583
           |: | :|: ||   : || |:  | ::|   |:| :     ||:| ||:|:|   |:| |
```

```
Db    503 AQAPKIWIYPTKTGSTYTEDSNPTITISGT-LTLRNSNNEDPYDSLDLSHSLEKVPLLYI 561

Qy    584 SEASDNQLRSDDMDFSGLNV-PHYGWQGLWTWGWAKTQDPEPASSATITDPQK---ANRF 639
          : :  :: | :| | ||   |||:||:|:  | :|         |||:|     ||
Db    562 VDVAAQKINSSQLDLSTLNSGEHYGYQGIWSTYWVET--------TTITNPTSLLGANTK 613

Qy    640 HRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSDHPFWGITGG-- 697
          |: |   | | || | |: |   | | || :   |   | : :         |    |
Db    614 HKLLYANWSPLGYRPHPERRGEFITNALWQSAYTALAGLHSLSS--------WDEEKGHA 665

Qy    698 ----GLGMMVYQDPRENHPGFHMRSSGYSAGMIA--GQTHTFSLKFSQTYTKLNERYAKN 751
              |:|::|:|  :    ||    :|||| |   |:  ||| |:| ::|  |  ::|
Db    666 ASLQGIGLLVHQKDKNGFKGFRSHMTGYSATTEATSSQSPNFSLGFAQFFSKAKEHESQN 725

Qy    752 NVSSKNY---SCQGEMLFSLQEGFLLTKLVGL-YSYGDHNCHHFYTQGENLTSQGTFRSQ 807
          : || :|    |    ||  :|   |:  | | | :   : |  |        |||:| :
Db    726 STSSHHYFSGMCIENTLF--KEWIRLS--VSLAYMFTSEHTHTMYQGLLEGNSQGSFHNH 781

Qy    808 TMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVP 867
          |: ||:      :| | : :   ||: || |  :|: | | | : | ||   || :| :|
Db    782 TLAGALSCVFLPQPHGES-LQIYPFITALAIRGNLAAFQESGDHAREFSLHRPLTDVSLP 840

Qy    868 IGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYK 927
          :|:: |: |  : |  |  |::|:  ||||:| : ::|| |:| |    :| : :|:  |
Db    841 VGIRASWKNHHRVPLVWLTEISYRSTLYRQDPELHSKLLISQGTWTTQATPVTYNALGIK 900

Qy    928 ISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
          :    |   :||   |   ||||  :|||   :||
Db    901 VKNTMQVFPKVTLSLDYSADISSSTLSHYLNVASRMRF 938


RESULT 4
PM16_CHLPN
ID   PM16_CHLPN     STANDARD;      PRT;   952 AA.
AC   Q9Z882; Q9JSE7; Q9K2A5;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable outer membrane protein pmp16 precursor (Polymorphic membrane
DE   protein 16).
GN   PMP16 OR CPN0467 OR CP0285.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CWL029;
RX   MEDLINE=99206606; PubMed=10192388;
RA   Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA   Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT   "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL   Nat. Genet. 21:385-389(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AR39;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).

CC   -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC       (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001631; AAD18609.1; ALT_INIT.
DR   EMBL; AE002190; AAF38142.1; -.
DR   EMBL; AP002546; BAA98673.1; ALT_INIT.
DR   TIGR; CP0285; -.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   InterPro; IPR003357; OMP.
DR   Pfam; PF02385; OMP; 1.
DR   Pfam; PF02415; DUF145; 1.
KW   Outer membrane; Signal; Multigene family; Complete proteome.
FT   SIGNAL        1     27       POTENTIAL.
FT   CHAIN        28    952       PROBABLE OUTER MEMBRANE PROTEIN PMP16.
FT   CONFLICT    125    125       G -> A (IN REF. 3).
SQ   SEQUENCE   952 AA;  104196 MW;  F1CB050DA6B74B5A CRC64;

  Query Match             20.9%;  Score 1060.5;  DB 1;  Length 952;
  Best Local Similarity   30.6%;  Pred. No. 4.7e-59;
  Matches  303;  Conservative  164;  Mismatches  447;  Indels   77;  Gaps   28;

Qy      6 FFFLIGNSLSGLAREVPSRIFLMPNSVPD------------PTKESLSNK------ISL 46
          | | :||  : :    |:   |  :|:               |  :|||         :
Db      8 FLFYLGNFTACMFGMTPAVYSLQTDSLEKFALERDEEFRTSFPLLDSLSTLTGFSPITTF 67

Qy     47 TGDTHNLT-NCYLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESG 105
           |: || : :  | | : |  ||    : | ||:  ::|      ::  :|:|||   :|
Db     68 VGNRHNSSQDIVLSNYKSIDNILLLWTSAGGAVSCNNFL--LSNVEDHAFFSKNLAIGTG 125

Qy    106 GAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQ-NLYIN 164
          |||          |    ||:|| :|          ||:      ||||    :  |:
Db    126 GAIACQG----ACTITKNRGPLIFFSNRGLN-------NASTGGETRGGAIACNGDFTIS 174

Qy    165 HNHDVVGFMKNFSYVRGGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSN 224
          |     |: |     |||:||     :  |::  || :      |||  ||      :
Db    175 QNQGTFYFVNNSVNNWGGALSTNGHCRIQSNRAPLLFFN------NTAPSGGGALRSENT 228

Qy    225 SFESNNCDLFFINNACCAGGAIFSPI-CSLTGNRGNIVFYNNRCFKNVETASSEASDGGA 283
          :   |   ::| ||    |||| : :  ::  | |:::| ||       : :|    |||
Db    229 TISDNTRPIYFKNNCGNNGGAIQTSVTVAIKNNSGSVIFNNNTALSG--SINSGNGSGGA 286

Qy    284 IKVTTRLDVTGNRGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYID 343
          |  || | :  | | | |::|      ||||    :|: ::|  || ||   | |||: :
Db    287 I-YTTNLSIDDNPGTILFNNNYCIRDGGAICTQFLTIKNSGHVYFTNN-QGNWGGALMLL 344

Qy    344 GTSNSKISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQN 403
          |   : |::  | |  | |   |             | |||    :|  : |||
Db    345 QDSTCLLFAEQGNIAFQNNEVFLTTFG----------RYNAIHCTPNS-NLQLGANKGYT 393

Qy    404 LIFYDPIEVSNAGVS-VSFNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGF 462
          |:|||  :   : : ||  |:  |:::|| | :  |  :: |  : :     | ||
Db    394 TAFFDPIEHQHPTTNPLIFNPNANHQGTILFSSAYIPEASDYENNFISSSKNTSELRNGV 453

Qy    463 LCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNSAS-NASITLKHIGLNLSSILK 521
          | ||| |     :||| ||:: ||: | ::    |    | |  : : :: :|| |||
Db    454 LSIEDRAGWQFYKFTQKGGILKLGHAASIATTANSETPSTSVGSQVIINNLAINLPSILA 513

Qy    522 SGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPML 581
          | : | ||: |  :|  :| |   | :||   |:|:::    ||:| ||:  | :  :|
Db    514 KG-KAPTLWIRPLQSSAPFTEDNNPTITLSG-PLTLLNEENRDPYDSIDLSEPLQNIHLL 571

Qy    582 SISEASDNQLRSDDMDFSGLN-VPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFH 640
          |:|: :   : :|:     ||   |||:||:|:  | :|      ::|:|   : ||  :
Db    572 SLSDVTARHINTDNFHPESLNATEHYGYQGIWSPYWVET--ITTTNNASI---ETANTLY 626
```

```
Qy     641 RTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSD--HPFWGITGGG 698
Db     627 RALYANWTPLGYKVNPEYQGDLATTPLWQSFHTMFSLLRSYNRTGDSDIERPFLEIQGIA 686

Qy     699 LGMMVYQDPRENHPGFHMRSSGYS--AGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSK 756
Db     687 DGLFVHQNSIPGAPGFRIQSTGYSLQASSETSLHQKISLGFAQFFTRTKEIGSSNNVSAH 746

Qy     757 N--YSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVF 814
Db     747 NTVSSLYVELPW-FQEAF-ATSTVLAYGYGDHHLHSLHPSHQE-QAEGTCYSHTLAAAIG 803

Qy     815 FDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSF 874
Db     804 CSFPWQQKSYLHL--SPFVQAIAIRSHQTAFEEIGDNPRKFVSQKPFYNLTLPLGIQGKW 861

Qy     875 MNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQP 934
Db     862 QSKFHVPTEWTLELSYQPVLYQQNPQIGVTLLASGGSWDILGHNYVRNALGYKVHNQTAL 921

Qy     935 LSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
Db     922 FRSLDLFLDYQGSVSSSTSTHHLQAGSTLKF 952
```

```
RESULT 5
PM18_CHLPN
ID   PM18_CHLPN     STANDARD;      PRT;   946 AA.
AC   Q9Z880; Q9RB60;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable outer membrane protein pmp18 precursor (Polymorphic membrane
DE   protein 18).
GN   PMP18 OR CPN0471 OR CP0283.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CWL029;
RX   MEDLINE=99206606; PubMed=10192388;
RA   Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA   Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT   "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL   Nat. Genet. 21:385-389(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AR39;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
CC   -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC       (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001631; AAD18610.1; -.
DR   EMBL; AE002189; AAF38141.1; -.
DR   EMBL; AP002546; BAA98677.1; -.
DR   TIGR; CP0283; -.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   InterPro; IPR003357; OMP.
DR   Pfam; PF02385; OMP; 1.
DR   Pfam; PF02415; DUF145; 1.
KW   Outer membrane; Signal; Multigene family; Complete proteome.
FT   SIGNAL        1     16       POTENTIAL.
FT   CHAIN        17    946       PROBABLE OUTER MEMBRANE PROTEIN PMP18.
FT   CONFLICT    426    426       A -> T (IN REF. 1).
FT   CONFLICT    882    882       V -> I (IN REF. 1).
SQ   SEQUENCE   946 AA;  103611 MW;  DCBF6B3D04AC452C CRC64;

  Query Match             18.7%;  Score 951.5;  DB 1;  Length 946;
  Best Local Similarity   29.1%;  Pred. No. 3.3e-52;
  Matches  281;  Conservative  168;  Mismatches  416;  Indels  101;  Gaps   34;

Qy      35 PTKESLSNKISLTGDTHNLT----------NCYLDNLRYILAILQKTPNEGAAVTITDYL 84
Db      47 PLIDTLTN---MTPYSHRATLFGVRDDTNQDIVLDHQNSIESWFENFSQDGGALSCKS-L 102

Qy      85 SFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNPN 144
Db     103 AITNT-KNQILFLNSFAIKRAGAM-YVNGN---FDLSENHGSIIFSGNL-------SFPN 150

Qy     145 AA--VNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSCFLFM 202
Db     151 ASNFADTCTGGAVLCSKNVTISKNQGTAYFINNKAKSSGGAIQAAIINIKDNTGPCLFF- 209

Qy     203 DNICIQTNTAG--KGGAIYAGTSNSFESNNCDLFFINNACCAGGAI-FSPICSLTGNRGN 259
Db     210 ------NNAAGGTAGGALFANACR-IENNSQPIYFLNNQSGLGGAIRVHQECILTKNTGS 262

Qy     260 IVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYAPVVT 319
Db     263 VIFNNN--FAMEADISANHSSGGAIYCIS-CSIKDNPGIAAFDNNTAARDGGAICTQSLT 319

Qy     320 LVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTSTSANP 379
Db     320 IQDSGPVYFTNN-QGTWGGAIMLRQDGACTLFADQGDIIFYNNRHFKDTFSN--HVSVNC 376

Qy     380 PRRNAITVASSSGEILLGAGSSQNLIFYDPI----EVSNAGVSVSFNKEADQTGSVVFSG 435
Db     377 TRNVSLTVGASQG---------HSATFYDPILQRYTIQNS--IQKFNPNPEHLGTILFSS 425

Qy     436 A-----TVNSADF--HQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNG 488
Db     426 AYIPDTSTSRDDFISHFRN-------HIGLYNGTLALEDRAEWKVYKFDQFGGTLRLGSR 478

Qy     489 AVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATF 548
Db     479 AVFSTTDEEQSSSSVGSVININNLAINLPSIL-GNRVAPKLWIRPTGSSAPYSEDNNPII 537

Qy     549 SLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSGLN-VPHYG 607
Db     538 NLSG-PLSLLDDENLDPYDTADLAQPIAEVPLLYLLDVTAKHINTDNFYPEGLNTTQHYG 596

Qy     608 WQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTL 667
Db     597 YQGVWSPYWIETITTSDTSS-----EDTVNTLHRQLYGDWTPTGYKVNPENKGDIALSAF 651

Qy     668 W---GNMLLATESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAG 724
Db     652 WQSFHNLFATLRYQTQQGQIAPT------ASGEATRLFVHQNSNNDAKGFHMEATGYSLG 705
```

```
Qy    725 MIA--GQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFS-LQEGFLLTKLVGLY 781
Db    706 TTSNTASNHSFGVNFSQLFSNLYESHSDNSVASHTTTVALQINNPWLQERFSTSASLA-Y 764

Qy    782 SYGDHNCHHFYTQGEN--LTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIY 839
Db    765 SYSN---HHIKASGYSGKIQTEGKCYSTTLGAALSCSLSLQ-WRSRPLHFTPFIQAIAVR 820

Qy    840 SSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEP 899
Db    821 SNQTAFQESGDKARKFSVHKPLYNLTVPLGIQSAWESKFRLPTYWNIELAYQPVLYQQNP 880

Qy    900 GIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNG 959
Db    881 EVNVSLESSGSSWLLSGTTLARNAIAFKGRNQIFIFPKLSVFLDYQGSVSSSTTTHYLHA 940

Qy    960 EIALRF 965
Db    941 GTTFKF 946
```

```
RESULT 6
PMPF_CHLTR
ID   PMPF_CHLTR     STANDARD;      PRT;  1034 AA.
AC   P38008; O84878;
DT   01-OCT-1994 (Rel. 30, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable outer membrane protein pmpF precursor (Polymorphic membrane
DE   protein F).
GN   PMPF OR CT870.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D/UW-3/Cx;
RX   MEDLINE=99000809; PubMed=9784136;
RA   Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA   Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA   Davis R.W.;
RT   "Genome sequence of an obligate intracellular pathogen of humans:
RT   Chlamydia trachomatis.";
RL   Science 282:754-759(1998).
RN   [2]
RP   SEQUENCE OF 26-35.
RC   STRAIN=L2/434/Bu;
RA   Bini L., Santucci A., Magi B., Marzocchi B., Sanchez-Campillo M.,
RA   Comanducci M., Christianen G., Birkelund S., Vtretou E., Ratti G.,
RA   Pallini V.;
RL   Submitted (SEP-1994) to the SWISS-PROT data bank.
CC   -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES).
CC   -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001360; AAC68468.1; -.
DR   Siena-2DPAGE; P38008; -.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   InterPro; IPR003357; OMP.
DR   Pfam; PF02385; OMP; 1.
DR   Pfam; PF02415; DUF145; 1.
KW   Outer membrane; Signal; Multigene family; Complete proteome.
FT   SIGNAL        1     25
FT   CHAIN        26   1034       PROBABLE OUTER MEMBRANE PROTEIN PMPF.
FT   CONFLICT     34     35       RR -> FH (IN REF. 2).
```

```
SQ   SEQUENCE   1034 AA;  112392 MW;  445FF4C35D463AE7 CRC64;

  Query Match            17.0%;  Score 866.5;  DB 1;  Length 1034;
  Best Local Similarity  28.1%;  Pred. No. 8.2e-47;
  Matches  293;  Conservative  149;  Mismatches  406;  Indels  193;  Gaps   36;

Qy     55 NCYLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAI------ 108
Db     57 NIVLSNLQ---------SNGTGACTISG-----NTQTQIFSNSVNTTADSGGAFDMVTTS 102

Qy    109 -------------GYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNPNAAV------NK 149
Db    103 FTASDNANLLFCNNYCTHNKGGGAIRSG-GPIRFLNNQDVLFYNNISAGAKYVGTGDHNE 161

Qy    150 IREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSCFLFMDNICI-- 207
Db    162 KNRGGALYATTITLTGNR-TLAFINNMSGDCGGAISADTQISITDTVKGILFENNHTLNH 220

Qy    208 ----QTNTAGKGGAIYAGTS-NSFESNNCDLFFINNACCAGGAIFSPICSLTGNRGNIVF 262
Db    221 IPYTQAENMARGGAICSRRDLCSISNNSGPIVFNYNQGGKGGAISATRCVIDNNKERIIF 280

Qy    263 YNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYAPVVTLVD 322
Db    281 SNNSSLG--WSQSSSASNGGAIQTTQGFTLRNNKGSIYFDSNTATHAGGAINCGYIDIRD 338

Qy    323 NGPTYFINNIANNKGGAIYIDGTSNSKISADRHA----------IIFNENIVTNVTNANG 372
Db    339 NGPVYFLNNSA--AWGAAF-------NLSKPRSATNYIHTGTGDIVFNNNVVFTLDGNLL 389

Qy    373 TSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIE-------VSNAGVS------V 419
Db    390 GKRKLFHINNNEITPYTLS----LGAKKDTRIYFYDLFQWERVKENTSNNPPSPTSRNTI 445

Qy    420 SFNKEADQTGSVVFSGATVNSADFHQRNLQTK-----TPAPLTLSNGFLCIEDHAQLTVN 474
Db    446 TVNPETEFSGAVVFS---YNQMSSDIRTLMGKEHNYIKEAPTTLKFGTLAIEDDAELEIF 502

Qy    475 R--FTQT-GGVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWV 531
Db    503 NIPFTQNPTSLLALGSGATLTVGKHG--------KLNITNLGVILPIILKEGKSPPCIRV 554

Qy    532 EPTN-NSNNYTADT-AATFSLS------DVKLSLIDDYGNSPYESTDLTHALSSQPMLSI 583
Db    555 NPQDMTQNTGTGQTPSSTSSISTPMIIFNGRLSIVDENYESVYDSMDLSRGKAEQLILSI 614

Qy    584 SEASDNQLRSD---DMDFSGLNVPHYGWQGLWTWGWAKT--------QDPEPASSATITD 632
Db    615 ETTNDGQLDSNWQSSLNTSLLSPPHYGYQGLWTPNWITTTYTITLNNNSSAPTSATSIAE 674

Qy    633 PQKANR----------------------------------FHRTLLLTWLPAGYVPSPKH 658
Db    675 QKKTSETFTPSNTTTASIPNIKASAGSGSGSASNSGEVTITKHTLVVNWAPVGYIVDPIR 734

Qy    659 RSPLIANTL---WGNMLLATESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFH 715
Db    735 RGDLIANSLVHSGRNMTMGLRSLL-------PDNSWFALQGAATTLFTKQQKRLSYHGYS 787

Qy    716 MRSSGYSAGMIA--GQTHTFSLKFSQTYTKLNERYAKNNVSSKNY---SCQGEMLFSLQE 770
Db    788 SASKGYTVSSQASGAHGHKFLLSFSQSSDKMKEKETNNRLSSRYYLSALCFEHPMFD--- 844

Qy    771 GFLLTKLVG--LYSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDL----PMKPFGS 824
Db    845 ---RIALIGAAACNYGTHNMRSFY--GTKKSSKGKFHSTTLGASLRCELRDSMPLR---- 895

Qy    825 THILTAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAW 884
Db    896 -SIMLTPFAQALFSRTEPASIRESGDLARLFTLEQAHTAVVSPIGIKGAYSSDTWPTLSW 954

Qy    885 TVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQY 944
Db    955 EMELAYQPTLYWKRPLLNTLLIQNNGSWVTTNTPLAKHSF-YGRGSHSLKFSHLKLFANY 1013
```

```
Qy     945 HGFYSSSTFCNYLNGEIALRF 965
Db    1014 QAEVATSTVSHYINAGGALVF 1034


RESULT 7
PMPF_CHLMU
ID   PMPF_CHLMU     STANDARD;      PRT;  1025 AA.
AC   Q9PL46;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable outer membrane protein pmpF precursor (Polymorphic membrane
DE   protein F).
GN   PMPF OR TC0262.
OS   Chlamydia muridarum.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=83560;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MoPn / Nigg;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
CC   -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC       (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE002293; AAF39131.1; -.
DR   TIGR; TC0262; -.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   InterPro; IPR003357; OMP.
DR   Pfam; PF02385; OMP; 1.
DR   Pfam; PF02415; DUF145; 1.
KW   Outer membrane; Signal; Multigene family; Complete proteome.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21   1025       PROBABLE OUTER MEMBRANE PROTEIN PMPF.
SQ   SEQUENCE   1025 AA;  111458 MW;  EA472E07818B9FF2 CRC64;

  Query Match             15.6%;  Score 793;  DB 1;  Length 1025;
  Best Local Similarity   27.0%;  Pred. No. 3.4e-42;
  Matches  269;  Conservative  158;  Mismatches  419;  Indels  152;  Gaps   35;

Qy      68 LQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPV 127
Db      80 LSTTNSSGAFV--ESMTSFTAIDNADLFFCNNYCTHQGG--GGAINATGLISFKNNQNIL 135

Qy     128 IFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTA 187
Db     136 FYNNTTIGTQFTGVALRTERNR---GGALYGSSIELINNHS-LNFINNTSGDMGGAVSTI 191

Qy     188 NTFVVSENQSCFLFMDNICIQ--TNT----AGKGGAIYAGTSNSFESNNCDLFFINNACC 241
Db     192 QNLVIKNTSGIVAFENNHTTDHIPNTFATILARGGAVGCQGACEISHNTGPVVFNSNYGG 251

Qy     242 AGGAIFS-PICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 300
Db     252 YGGAISTGGQCIFRDNKDKLIFINNSALGWHNT--SAQGNGAVISAGGEFGLLNNKGPIY 309

Qy     301 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGT-SNSKISADRHAIIF 359
Db     310 FENNNASYIAGAISCNNLNFQENGPIYFLNNSA-LYGGAFHLFASPAANYIHTGSGDIIF 368

Qy     360 NENIVTNVTNAN----------GTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDP 409
Db     369 NNNTELSTTGMSAGLRKLFYIPGTTNN------NPITLS-------LGAKKDTRIYFYDL 415

Qy     410 IE----------VSNAGVSVSFNKEADQTGSVVFSGATVNSADFHQRNLQTKT-----PA 454
Db     416 FQWGGLKKANTPPENSPHTVTINPSDEFSGAVVFSYKNI-SSDLQAHMIASKTHNQIKDS 474

Qy     455 PLTLSNGFLCIEDHAQ-------LTVNRFTQTGGVVSLGNGAVLSCYKNGAGNSASNASI 507
Db     475 PTTLKFGTMSIENGAEFEFFNGPLT----QESTSLLALGQDSILTVGK--------DASL 522

Qy     508 TLKHIGLNLSSILK-SGAEIPLLWVEP---TNNSNNYTADTAATFSLSDVKL------SL 557
Db     523 TITHLGIILPGLLNDQGTTAPRIRVNPQDMTQNTNSNQA-PVSTENVATQKIFFSGLVSL 581

Qy     558 IDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSD---DMDFSGLNVPHYGWQGLWTW 614
Db     582 VDENYESVYDSCDLSRGKANQPILHIETTNDAQLSNDWKNTLNTSLYSLPHYGYQGLWTS 641

Qy     615 GWAKTQDPEPASSATITDPQKANR---------------------------------- 638
Db     642 NWMTT--TRTVSLTNSTETQTANNSIQEQKNTSETFDSNSTTTAKIPSIRASTGGTTPLA 699

Qy     639 ------FHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSDHPFW 692
Db     700 TTDVTVTRHSLVVSWTPIGYIADPARRGDLIANNLVSSGRNTTLYLRS---LLPDD-SWF 755

Qy     693 GITGGGLGMMVYQDPRENHPGFHMRSSGY--SAGMIAGQTHTFSLKFSQTYTKLNERYAK 750
Db     756 ALQGSAATLFTKQQKRLDYHGYSSASKGYAISSQASGAHGHKFLFSFSQSSDTMKEKRTN 815

Qy     751 NNVSSKNYSCQGEMLFSLQEGFLLTKLVG--LYSYGDHNCHHFYTQGENLTSQGTFRSQT 808
Db     816 NKISSRYYL---SALCFEQPMFDRIALIGAAAYNYGTHKTYNFY--GTKKFSKGNFHSTT 870

Qy     809 MGGAVFFDL-PMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVP 867
Db     871 LGGSLRCELRDSMPFQS--IMLTPFIQALISRTEPASIQEQGDLARLFSLKQPHTAVVSP 928

Qy     868 IGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYK 927
Db     929 IGIKGVYSSNKWPTVSCEMEVAYQPTLYWKRPILNTVLIKNNGSWETTNTPLAKHSF-YG 987

Qy     928 ISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
Db     988 RGSSSLKFSYLKLFANYQAQVATSTVSHYMNAGGALVF 1025


RESULT 8
PMPG_CHLTR
ID   PMPG_CHLTR     STANDARD;      PRT;  1013 AA.
AC   O84879;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable outer membrane protein pmpG precursor (Polymorphic membrane
DE   protein G).
GN   PMPG OR CT871.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D/UW-3/Cx;
RX   MEDLINE=99000809; PubMed=9784136;
RA   Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA   Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
```

```
RA   Davis R.W.;
RT   "Genome sequence of an obligate intracellular pathogen of humans:
RT   Chlamydia trachomatis.";
RL   Science 282:754-759(1998).
CC   -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC       (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001360; AAC68469.1; -.
DR   PHCI-2DPAGE; O84879; -.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   InterPro; IPR003357; OMP.
DR   Pfam; PF02385; OMP; 1.
DR   Pfam; PF02415; DUF145; 1.
KW   Outer membrane; Signal; Multigene family; Complete proteome.
FT   SIGNAL        1     27       POTENTIAL.
FT   CHAIN        28   1013       PROBABLE OUTER MEMBRANE PROTEIN PMPG.
SQ   SEQUENCE   1013 AA;  107366 MW;  F0927743C0A651DD CRC64;

  Query Match             12.9%;  Score 655.5;  DB 1;  Length 1013;
  Best Local Similarity   25.7%;  Pred. No. 1.5e-33;
  Matches  261;  Conservative  152;  Mismatches  388;  Indels  213;  Gaps   43;

Qy      39 SLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAK 98
           :|    :: |  |:||    :|:|        |   |||::        |:   |::
Db      84 NLLGSFTVLGRGHSLT---FENIR--------TSTNGAALS--------DSANSGLF--- 121

Qy      99 NLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNP-NAAV---------- 147
             | |    : ::: ||    :     |    ||    | |:| | |  :
Db     122 --TIEGFKELSFSNCNSLLAVL-----PAATTNNGSQTPTTTSTPSNGTIYSKTDLLLLN 174

Qy     148 --------NKIR-EGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSC 198
                    | :  :|||| |::| :     :  | :| :   |||     :|    |::
Db     175 NEKFSFYSNLVSGDGGAIDAKSLTVQGISKLCVFQENTAQADGGACQVVTSFSAMANEAP 234

Qy     199 FLFMDNICIQTNTAG-KGGAIYA------GTSNSFE--------SNNCDLFFINNACCAG 243
             |: |:      || :|| | |      | |:|          | |  : |  |    |
Db     235 IAFIANV------AGVRGGGIAAVQDGQQGVSSSTSTEDPVVSFSRNTAVEFDGNVARVG 288

Qy     244 GAIFSPICSLTGNRGNIVFYNN--RCFKN----------------VETASSEASDGGAI 284
           | |:|          ||: | ||    | |                    |||||
Db     289 GGIYS--------YGNVAFLNNGKTLFLNNVASPVYIAAEQPTNGQASNTSDNYGDGGAI 340

Qy     285 KVTTRLDVTGN----------RGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIAN 334
                  |:           | :||| |:    ||||||  ::: : ||  |: ||||
Db     341 FCKNGAQAAGSNNSGSVSFDGEGVVFFSSNVAAGKGGAIYAKKLSVANCGPVQFLGNIAN 400

Qy     335 NKGGAIYIDGTSNSKISADRHAIIFNENI----VTNVTNANGTSTSANPPRRNAITVASS 390
           : |||||:  :    :|||   |||: |:      |  : || : |:       ||:: |
Db     401 D-GGAIYLGESGELSLSADYGDIIFDGNLKRTAKENAADVNGVTVSS-----QAISMGSG 454

Qy     391 SGEILLGAGSSQNLIFYDPIEVSN-------AGVSVSFNKEADQTGSVVFSGATVNSADF 443
             | | :   ::| ||||::|       :   :  |     || :||  |  ||  :
Db     455 GKITTLRAKAGHQILFNDPIEMANGNNQPAQSSEPLKINDGEGYTGDIVF--ANGNSTLY 512

Qy     444 HQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNS-- 501
                     :|:  | : : : |:|:||  :|||| : :  |: |
Db     513 QN----------VTIEQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTLDFVTPQPPQQPP 562

Qy     502 ASNASITLKHIGLNLSSILKSGAEIPLLWVEPTN--NSNNYTADTAATFSLSDVKLS--- 556
           |:|  ||| :: |:|||:| : |    :   |||    ::: |  :| :   | :|
Db     563 AANQLITLSNLHLSLSSLLANNA----VTNPPTNPPAQDSHPAIIGST-TAGSVTISGPI 617

Qy     557 LIDDYGNSPYESTD----------LTHALSSQPMLSISEASDNQLRSDDMDFSGLNVPHY 606
             :|  :: |:  |          |   | :||  | :  ||  |        |  :| |
Db     618 FFEDLDDTAYDRYDWLGSNQKIDVLKLQLGTQP--SANAPSDLTL--------GNEMPKY 667

Qy     607 GWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANT 666
           |:|| |   |               ||  ||    ||  ||   || | |:  : |: |:
Db     668 GYQGSWKLAW--------------DPNTANNGPYTLKATWTKTGYNPGPERVASLVPNS 712

Qy     667 LWGNMLLATESLKNSAELTPSDHPFW-GITGGGLGMMVYQDPRENHPGFHMRSSGYSAGM 725
           |||: :|   |  :: : :     :  |:   |:    | |      |:   | ||| |
Db     713 LWGS-ILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYHDRDALGQGYRYISGGYSLG- 770

Qy     726 IAGQTHTFSLKFSQTYTKLNERYAKNNV--SSKNYSCQGEMLFSLQEGFLLTKLVGLYSY 783
               ::  |  |   :|::  | :|: |   | :::| | :  | ::      | | | :
Db     771 --ANSYFGSSMFGLAFTEVFGR-SKDYVVCRSNHHACIGSVYLSTKQA-----LCGSYLF 822

Qy     784 GDHNCHHFYTQG-ENLTSQGTFRSQT--------MGGAVFFDLPMKPFGSTHILT--APF 832
           ||       |  | ::: :  ||  ::         : | :   ||:    |   |    ||
Db     823 GDAFIRASYGFGNQHMKTSYTFAEESDVRWDNNCLVGEIGVGLPIVITPSKLYLNELRPF 882

Qy     833 LGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQP 892
           : |   |:    ||| |   |:| :   |:|: ||:|||    ::|  |  ::   ||
Db     883 VQAEFSYADHESFTEEGDQARAFRS-GHLMNLSVPVGVKFDRCSST-HPNKYSFMGAYIC 940

Qy     893 VLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHG 946
             ||   |  | ||: :  |       :|| :  : |        || : : :|
Db     941 DAYRTISGTQTTLLSHQETWTTDAFHLARHGVIVRGSMYAS----LTSNIEVYG 990


RESULT 9
PMP7_CHLPN
ID   PMP7_CHLPN     STANDARD;      PRT;   936 AA.
AC   Q9Z898; Q9JS42; Q9Z5Q4;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable outer membrane protein pmp7 precursor (Polymorphic membrane
DE   protein 7) (Outer membrane protein 12).
GN   PMP7 OR OMP12 OR CPN0445 OR CP0308.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CWL029;
RX   MEDLINE=99206606; PubMed=10192388;
RA   Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA   Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT   "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL   Nat. Genet. 21:385-389(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AR39;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
RN   [4]
RP   SEQUENCE OF 658-936 FROM N.A.
RC   STRAIN=VR1310;
RX   MEDLINE=20007584; PubMed=10539856;
```

```
RA   Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
RA   Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT   "Molecular biology of Chlamydia pneumoniae surface proteins and their
RT   role in immunopathogenicity.";
RL   Am. Heart J. 138:S491-S495(1999).
CC   -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC       (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001627; AAD18589.1; -.
DR   EMBL; AE002193; AAF38165.1; -.
DR   EMBL; AP002546; BAA98653.1; -.
DR   EMBL; AJ133034; CAB37067.1; -.
DR   PHCI-2DPAGE; Q9Z898; -.
DR   TIGR; CP0308; -.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   InterPro; IPR003357; OMP.
DR   Pfam; PF02385; OMP; 1.
DR   Pfam; PF02415; DUF145; 2.
KW   Outer membrane; Signal; Multigene family; Complete proteome.
FT   SIGNAL       1     23       POTENTIAL.
FT   CHAIN       24    936       PROBABLE OUTER MEMBRANE PROTEIN PMP7.
FT   CONFLICT   658    666       PTRHGFRHI -> EDNIRYRHN (IN REF. 4).
FT   CONFLICT   822    822       Y -> H (IN REF. 1 AND 4).
SQ   SEQUENCE   936 AA;  100105 MW;  3981DB3C950AF95A CRC64;

  Query Match           12.7%;  Score 644;  DB 1;  Length 936;
  Best Local Similarity  27.0%;  Pred. No. 6.9e-33;
  Matches  262;  Conservative  146;  Mismatches  392;  Indels  172;  Gaps   44;

Qy     75 GAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGY-------------ASPNSPTVEIR 121
          |   ::   :|| :    ||  |     |:|| : :             |  :: ||
Db     54 GTTYSLLSDVSFQNAGALGIPLASGCFLEAGGDLTFQGNQHALKFAFINAGSSAGTVAST 113

Qy    122 DTIGPVIFENN-------TCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMK 174
            :  |:       :|     |     |:  :          || :  |  :: | :
Db    114 SAADKNLLFNDFSRLSIISCPSLLLSPTGQCALKSV--------GNLSLTGNSQII-FTQ 164

Qy    175 NFSYVRGGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYA-GTSNSFESNNCDL 233
          |||   || |:| | |::|       |  |   |  |  :|| :|| ||     |
Db    165 NFSSDNGGVINTKN-FLLSGTSQFASFSRN---QAFTGKQGGVVYATGTITIENSPGIVS 220

Qy    234 FFINNACCAGGAIFS-PICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTT---R 289
          |  | |  :|||::|   ||:| |   ::|  |  ::  :      : ||||  ||
Db    221 FSQNLAKGSGGALYSTDNCSITDN-FQVIFDGNSAWEAAQ------AQGGAICCTTTDKT 273

Qy    290 LDVTGNRGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNK-----GGAIYIDG 344
          : :|||:  : |::|    |||||    |::   ||| | :||: :      ||||
Db    274 VTLTGNK-NLSFTNNTALTYGGAISGLKVSISAGGPTLFQSNISGSSAGQGGGGAINIAS 332

Qy    345 TSNSKISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNL 404
              :||    | || | |||      |||      |||| :  ::    : | : |::
Db    333 AGELALSATSGDITFNNNQVTN----GSTST------RNAINIIDTAKVTSIRAATGQSI 382

Qy    405 IFYDPIEVSNAGVSVS----------FNKEADQTGSVVFSGATVNSAD-FHQRNLQTKTP 453
           |||||  :| | : |           | | :  |::|||  ::  :      |: :
Db    383 YFYDPI--TNPGTAASTDTLNLNLADANSEIEYGGAIVFSGEKLSPTEKAIAANVTSTIR 440

Qy    454 APLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNSASNASITLKHIG 513
           |  |: | | : |   :|    ||: |   | :|        |   ||  |:::|  :
Db    441 QPAVLARGDLVLRDGVTVTFKDLTQSPGSRILMDG--------GTTLSAKEANLSLNGLA 492

Qy    514 LNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTH 573
          :||||:           : || :   |       |||   ::|||  | | ||: :|
Db    493 VNLSSL-----------DGTNKAALKTEAADKNISLSGT-IALIDTEG-SFYENHNLKS 538

Qy    574 ALSSQPMLSISEASDN-QLRSDDMDFSGLNVP--HYGWQGLWTWGWAKTQDPEPASSATI 630
          | |: |:| :: |  |  :    :    |  |  |||:|| |   ||       |:|: |
Db    539 A-STYPLLELTTAGANGTITLGALSTLTLQEPETHYGYQGNWQLSWAN------ATSSKI 591

Qy    631 TDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSDHP 690
                      : |   ||:|||: :| |  |:||||  :   |:    |   |  |
Db    592 GS-----------INWTRTGYIPSPERKSNLPLNSLWGN-FIDIRSINQLIETKSSGEP 638

Qy    691 F----WGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFSLKFSQTYTKLNE 746
          |    |     |:   |:|      ||   | ||: |: |       | |:  : :|
Db    639 FERELW---LSGIANFFYRDSMPTRHGFRHISGGYALGITATTPAEDQLTFA--FCQLFA 693

Qy    747 RYAKNNVSSKN---------YSCQGEMLFSLQEGFL-------------LTKLVGL---- 780
          |  :|::: ||          |    | || :   ||             :::::  |
Db    694 R-DRNHITGKNHGDTYGASLYFHHTEGLFDI-ANFLWGKATRAPWVLSEISQIIPLSFDA 751

Qy    781 ---YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHIL--TAPFLGA 835
             | : |::   :|| :|   :|::|:      :   ||       ::|    ||:
Db    752 KFSYLHTDNHMKTYYT--DNSIIKGSWRNDAFCADLGASLPF-VISVPYLLKEVEPFVKV 808

Qy    836 LGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVL- 894
            ||:    | |  |  |:|: |: |||| :||||         : :  | :|    :|
Db    809 QYIYAHQQDFYERYAEGRAFN-KSELINVEIPIGVT---FERDSKSEKGTYDLTLMYILD 864

Qy    895 -YRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTF 953
             ||: |   | |:||   |   |:  :|   | : :  |    : :  |:  |   |:
Db    865 AYRRNPKCQTSLIASDANWMAYGTNLARQGFSVRAANHFQVNPHMEIFGQF-AFEVRSSS 923

Qy    954 CNYLNGEIALRF 965
             || |  :  :|
Db    924 RNY-NTNLGSKF 934


RESULT 10
PMPG_CHLMU
ID   PMPG_CHLMU     STANDARD;      PRT;   987 AA.
AC   Q9PL45;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable outer membrane protein pmpG precursor (Polymorphic membrane
DE   protein G).
GN   PMPG OR TC0263.
OS   Chlamydia muridarum.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=83560;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MoPn / Nigg;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
CC   -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC       (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE002293; AAF39132.1; -.
DR   TIGR; TC0263; -.
```

```
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   InterPro; IPR003357; OMP.
DR   Pfam; PF02385; OMP; 1.
DR   Pfam; PF02415; DUF145; 1.
KW   Outer membrane; Signal; Multigene family; Complete proteome.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26    987       PROBABLE OUTER MEMBRANE PROTEIN PMPG.
SQ   SEQUENCE   987 AA;  104867 MW;  32079BD6BEB2DA42 CRC64;

  Query Match             12.2%;  Score 620.5;  DB 1;  Length 987;
  Best Local Similarity   25.8%;  Pred. No. 2.2e-31;
  Matches  253;  Conservative  153;  Mismatches  404;  Indels  171;  Gaps   42;

Qy      39 SLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEGAA-----------VTITDYLSFF 87
           :|    :: |  |:|     :|:|        |   |||           :   | ||
Db      82 NLLGNFTIAGRGHSLV---FENIR--------TSTNGAALSNHAPSGLFVIEAFDELSLL 130

Qy      88 DTQKEGIYFAKNLTPESGGAIGYASPNSPTV------EIRDTIGPVIFENNTCCRPFTSS 141
           :          :: |::||    : |:: |:        :|| |  | | :|
Db     131 NCNS-----LVSVVPQTGGTT-TSVPSNGTIYSRTDLVLRD-IKKVSFYSNLVS------ 177

Qy     142 NPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSCFLF 201
                  :|||| ||:| :|    :  | :| :   |||     ||    |:    |
Db     178 ---------GDGGAIDAQSLMVNGIEKLCTFQENVAQSDGGACQVTKTFSAVGNKVPLSF 228

Qy     202 MDNICIQTNTAGKGGAIYA-----GTSNSFE-----SNNCDLFFINNACCAGGAIFSPIC 251
           : |:        ||| : |     |   : :     :||  : |  |:   || |:|
Db     229 LGNV-----AGNKGGGVAAVKDGQGAGGATDLSVNFANNTAVEFEGNSARIGGGIYSD-- 281

Qy     252 SLTGNRGNIVFYNNR---CFKNVETA--------------SSEASDGGAI--KVTTRLD 291
                |||  |   |        || :                  |||||  |  | :
Db     282 ------GNISFLGNAKTVFLSNVASPIYVDPAAAGGQPPADKDNYGDGGAIFCKNDTNIG 335

Qy     292 VTG--NRGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSK 349
               : | :||| ||    ||||||  :|: | ||   |: |:||: |||||:
Db     336 EVSFKDEGVVFFSKNIAAGKGGAIYAKKLTISDCGPVQFLGNVAND-GGAIYLVDQGELS 394

Qy     350 ISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDP 409
           :|||| |||: |:    |    |          |||::|:       | |   : ::| ||
Db     395 LSADRGDIIFDGNLKRMATQGAATVHDV-MVASNAISMATGGQITTLRAKEGRRILFNDP 453

Qy     410 IEVSNAG---VSVSFNKEADQTGSVVFS-GATVNSADFHQRNLQTKTPAPLTLSNGFLCI 465
           ||::|      ::: |:    || :||: |  |                : : || | : :
Db     454 IEMANGQPVIQTLTVNEGEGYTGDIVFAKGDNV-------------LYSSIELSQGRIIL 500

Qy     466 EDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSILK-SGA 524
             :| ||  ||||| | :  |: |           :  |: | ::  :|:|:|| :|
Db     501 REQTKLLVNSLTQTGGSVHMEGGSTLDF---AVTTPPAANSMALTNVHFSLASLLKNNGV 557

Qy     525 EIPLLWVEPTN-----NSNNYTADTAA-TFSLSDVKLSLIDDYGNSPYESTDLTHALSSQ 578
             |     |||      :|     :||| | ::|       :|   : |::     |  :
Db     558 TNP-----PTNPPVQVSSPAVIGNTAAGTVTISGP--IFFEDLDETAYDNNQWLGADQTI 610

Qy     579 PMLSISEASDNQLRSDDMDFSGLNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANR 638
           :| :   ::    :       |     ||:|| ||  |    :|:||      :| : |
Db     611 DVLQLHLGANPPANAPTDLTLGNESSKYGYQGSWTLQW----EPDPA------NPPQNNS 660

Qy     639 FHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSDHPFW-GITGG 697
           :   |  :|   || | |:  : |:::|:|||: :|   |  :: : :     :  ||
Db     661 Y--MLKASWTKTGYNPGPERVASLVSNSLWGS-ILDVRSAHSAIQASIDGRAYCRGIWIS 717

Qy     698 GLGMMVYQDPRENHPGFHMRSSGYSAGMIA--GQTHTFSLKFSQTYTKLNERYAKNNV-- 753
           |:    | |       |:   | |||  |  :  | :  | | |::|: :     :|: |
Db     718 GISNFFYHDQDALGQGYRHISGGYSIGANSYFGSS-MFGLAFTETFGR-----SKDYVVC 771

Qy     754 SSKNYSCQGEMLFSLQEGFLLTKLVG------LYSYGDHNCHHFYTQGENLTSQGTFRSQ 807
           | :::| | :  | ::    : | |       | :|: :    ||  |   |   : :
Db     772 RSNDHTCVGSVYLSTRQALCGSCLFGDAFVRASYGFGNQHMKTSYTFAEE--SNVRWDNN 829

Qy     808 TMGGAVFFDLPMKPFGSTHILT--APFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVL 865
           : | |   ||:    |   |    ||: |   |:    ||| |   | | :   |:|:
Db     830 CVVGEVGAGLPIMLAASKLYLNELRPFVQAEFAYAEHESFTERGDQAREFKS-GHLMNLS 888

Qy     866 VPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRHAMS 925
           :|:|||     ::: |  ::   ||    ||   |  | ||: |  |       :|| :
Db     889 IPVGVKFD-RCSSKHPNKYSFMGAYICDAYRSISGTETTLLSHKETWTTDAFHLARHGVM 947

Qy     926 YKISQQTQPLSWLTLHFQYHG 946
             : |        || : : :|
Db     948 VRGSMYAS----LTGNIEVYG 964


RESULT 11
PMP8_CHLPN
ID   PMP8_CHLPN     STANDARD;      PRT;   930 AA.
AC   Q9Z393; Q9RB66;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable outer membrane protein pmp8 precursor (Polymorphic membrane
DE   protein 8) (Outer membrane protein 11).
GN   PMP8 OR OMP11 OR CPN0446 OR CP0307.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VR1310;
RX   MEDLINE=20007584; PubMed=10539856;
RA   Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
RA   Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT   "Molecular biology of Chlamydia pneumoniae surface proteins and their
RT   role in immunopathogenicity.";
RL   Am. Heart J. 138:S491-S495(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CWL029;
RX   MEDLINE=99206606; PubMed=10192388;
RA   Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA   Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT   "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL   Nat. Genet. 21:385-389(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AR39;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
CC   -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC       (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ133034; CAB37068.1; -.
```

```
DR   EMBL; AE001627; AAD18590.1; -.
DR   EMBL; AE002193; AAF38164.1; -.
DR   EMBL; AP002546; BAA98654.1; -.
DR   PHCI-2DPAGE; Q9Z393; -.
DR   TIGR; CP0307; -.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   InterPro; IPR003357; OMP.
DR   Pfam; PF02385; OMP; 1.
DR   Pfam; PF02415; DUF145; 2.
KW   Outer membrane; Signal; Multigene family; Complete proteome.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27    930       PROBABLE OUTER MEMBRANE PROTEIN PMP8.
FT   CONFLICT    177    177       T -> A (IN REF. 3 AND 4).
SQ   SEQUENCE   930 AA;  97669 MW;  46A9B5E3BB913C4C CRC64;

  Query Match             12.1%;  Score 617;  DB 1;  Length 930;
  Best Local Similarity   27.0%;  Pred. No. 3.4e-31;
  Matches  251;  Conservative  140;  Mismatches  374;  Indels  164;  Gaps   40;

Qy      47 TGDTHNLTNCYLDNLRYILAILQKTPNEGAAVT---ITDYLSFFDTQKEGIYFAKNLTPE 103
           | |  | ||  |    ||         :| |:|    |:         :|  |: | | :
Db      50 TADA-NGTNYVLSGNVYI-----NDAGKGTALTGCCFTETTGDLTFTGKGYSFSFN-TVD 102

Qy     104 SGGAIGYASPNSPTVEIRDT----IGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQ 159
           :|   | |:  :    : |    :  :     |     :: :   |:|    |  : :|
Db     103 AGSNAGAAASTTADKALTFTGFSNLSFIAAPGTTVASGKSTLSSAGALNLTDNGTILFSQ 162

Qy     160 NLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSCFLFMDNICIQTNTAGK-GGAI 218
           |:    |::            ||||:|  |  :| | |      :|   :|:| | ||||
Db     163 NVSNEANNN-----------GGAITT-KTLSISGNTS------SITFTSNSAKKLGGAI 203

Qy     219 YAGTSNSFESNNCDLFFINN-ACCAGGAI-FSPICSLTGNRGNIVFYNNRCFKNVETASS 276
           |:  : |   |   | |:||     |||: |    |:| |  :: |  |       ||:
Db     204 YSSAAASISGNTGQLVFMNNKGETGGGALGFEASSSITQN-SSLFFSGN-------TATD 255

Qy     277 EASDGGAIKV-----TTRLDVTGNRGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINN 331
           |  ||||       |  | ::||:  : |::| :   ||| |  : |   ||| | ||
Db     256 AAGKGGAIYCEKTGETPTLTISGNKS-LTFAENSSVTQGGAICAHGLDLSAAGPTLFSNN 314

Qy     332 IANN----KGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITV 387
           |    ||||| |  : :  :||::  | |  | :        |||||    |||| :
Db     315 RCGNTAAGKGGAIAIADSGSLSLSANQGDITFLGNTL--------TSTSAPTSTRNAIYL 366

Qy     388 ASSSGEILLGAGSSQNLIFYDPIEVSNAGVS-------VSFNKEADQTGSVVFSGATVNS 440
           ||:    | |   |:: |||||  :  | |          |   | :|::|||  : |
Db     367 GSSAKITNLRAAQGQSIYFYDPIASNTTGASDVLTINQPDSNSPLDYSGTIVFSGEKL-S 425

Qy     441 ADFHQR--NLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGA 498
           ||  :   |  :    || |::| | :: : :| || |||| |   |       : |
Db     426 ADEAKAADNFTSILKQPLALASGTLALKGNVELDVNGFTQTEGSTLL--------MQPGT 477

Qy     499 GNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLI 558
           |   :|:|  : ::||::            |   : :  ||    | :|:   |
Db     478 KLKADTEAISLTKLVVDLSAL-----------EGNKSVSIETAGANKTITLTS-PLVFQ 524

Qy     559 DDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSGLNV-------PHYGWQGL 611
           |  ||  |||  :  |  :||::  : |:     : |:    |          ||||:||
Db     525 DSSGNF-YESHTINQAF-TQPLVVFTAAT----AASDIYIDALLTSPVQTPEPHYGYQGH 578

Qy     612 WTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNM 671
           |   ||   |   | | |                 :||:  || |:|: |: :: ::|| :
Db     579 WEATWA---DTSTAKSGT---------------MTWVTTGYNPNPERRASVVPDSLWASF 620

Qy     672 LLATESLKNSAELTPSDHPFW---GITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAG 728
           |:       :|   :  :   |:   |     ::|    :  |  :| ||  |  |
Db     621 ---TDIRTLQQIMTSQANSIYQQRGLWASGTANFFHKDKSGTNQAFRHKSYGYIVGGSAE 677

Qy     729 --QTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQ----------------GEMLFSLQ 769
             : ||: | | : |  : :   | :| ||                    | :   |:
Db     678 DFSENIFSVAFCQLFGKDKDLFIVEN-TSHNYLASLYLQHRAFLGGLPMPSFGSITDMLK 736

Qy     770 EGFLLTKLVGLYSYGDHNCHHFYT-----QGENLTSQGTFRSQTMGGAVFFDLPMK-PFG 823
           :  |:      |||  ::    ||     ||    : |      :||::   || : ||
Db     737 DIPLILNAQLSYSYTKNDMDTRYTSYPEAQGSWTNNSGALE---LGGSLALYLPKEAPFF 793

Qy     824 STHILTAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQA 883
             :    |||    :||   :| | ||  |:|     |:|  :|:|::   ::  ::
Db     794 QGYF---PFLKFQAVYSRQQNFKESGAEARAFD-DGDLVNCSIPVGIRLEKISEDEK-NN 848

Qy     884 WTVELAYQPVLYRQEPGIATQLLASKGIW 912
           : : |||   :||: |   | |: |   |
Db     849 FEISLAYIGDVYRKNPRSRTSLMVSGASW 877


RESULT 12
PM11_CHLPN
ID   PM11_CHLPN     STANDARD;      PRT;   928 AA.
AC   O86164; Q9K299;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable outer membrane protein pmp11 precursor (Polymorphic membrane
DE   protein 11) (Outer membrane protein 4).
GN   PMP11 OR OMP4 OR CPN0449 OR CP0302.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CWL029/VR-1310;
RX   MEDLINE=99081766; PubMed=9864239;
RA   Knudsen K., Madsen A.S., Mygind P., Christiansen G., Birkelund S.;
RT   "Identification of two novel genes encoding 97- to 99-kilodalton outer
RT   membrane proteins of Chlamydia pneumoniae.";
RL   Infect. Immun. 67:375-383(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VR1310;
RX   MEDLINE=20007584; PubMed=10539856;
RA   Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
RA   Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT   "Molecular biology of Chlamydia pneumoniae surface proteins and their
RT   role in immunopathogenicity.";
RL   Am. Heart J. 138:S491-S495(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CWL029;
RX   MEDLINE=99206606; PubMed=10192388;
RA   Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA   Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT   "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL   Nat. Genet. 21:385-389(1999).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AR39;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
CC   -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC       (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AJ001311; CAA04672.1; -.
DR   EMBL; AJ133034; CAB37072.1; -.
DR   EMBL; AE001628; AAD18593.1; -.
DR   EMBL; AE002192; AAF38159.1; ALT_INIT.
DR   EMBL; AP002546; BAA98658.1; -.
DR   PHCI-2DPAGE; O86164; -.
DR   TIGR; CP0302; -.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   InterPro; IPR003357; OMP.
DR   Pfam; PF02385; OMP; 1.
DR   Pfam; PF02415; DUF145; 2.
KW   Outer membrane; Signal; Multigene family; Complete proteome.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    928       PROBABLE OUTER MEMBRANE PROTEIN PMP11.
SQ   SEQUENCE   928 AA;  98903 MW;  788BCDD62C911402 CRC64;

  Query Match             11.7%;  Score 597;  DB 1;  Length 928;
  Best Local Similarity   26.1%;  Pred. No. 6.2e-30;
  Matches  253;  Conservative  148;  Mismatches  361;  Indels  206;  Gaps   47;

Qy      36 TKESLSNKISLTGD---------THNLTNCY---LDNLRYI-------LAILQKTPNEGA 76
           | :: :   ||||         |    :|:    ||| ::          :    : ||
Db      45 TPKTSATTYSLTGDVFFYEPGKGTPLSDSCFKQTTDNLTFLGNGHSLTFGFIDAGTHAGA 104

Qy      77 AVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCR 136
           | : |                ||||      : : |  | ||    | |
Db     105 AASTT-------------ANKNLTFSGFSLLSFDSSPSTTV----TTGQ---------- 136

Qy     137 PFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQ 196
                     :   | :: :|:       | |   |||  ||||  | :|:::
Db     137 -----------GTLSSAGGVNLENI---RKLVVAG---NFSTADGGAIKGA-SFLLTGTS 178

Qy     197 SCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDL--FFINNACCAGGAIFSPICSLT 254
           || :|     ::: |||||  | |: :  :||     |  | |  :||||     |:
Db     179 GDALFSNN-----SSSTKGGAI-ATTAGARIANNTGYVRFLSNIASTSGGAIDDEGTSIL 232

Qy     255 GNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLD----VTGNRGRIFFSDNITKNYG 310
           |   : |  |          :  : ||||  |         :  |   : |: |: :  |
Db     233 SNNKFLYFEGN----------AAKTTGGAICNTKASGSPELIISNNKTLIFASNVAETSG 282

Qy     311 GAIYAPVVTLVDNGPTYFI-NNI--ANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNV 367
           |||:|  : |   | | |: ||:  |  ||||| || :    :||:   | |  | :|
Db     283 GAIHAKKLALSSGGFTEFLRNNVSSATPKGGAISIDASGELSLSAETGNITFVRNTLT-- 340

Qy     368 TNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPI--EVSNAGVSVSFNKEA 425
              |: | : |:|||| : |:     | |  :  : |||||  | ::: |    |  |
Db     341 -----TTGSTDTPKRNAINIGSNGKFTELRAAKNHTIFFYDPITSEGTSSDVLKINNGSA 395

Qy     426 ----DQTGSVVFSGATVNSADFH-QRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQ-T 479
                |:::||| |: : :     ||::    |::|| | | ::    |    |:|
Db     396 GALNPYQGTILFSGETLTADELKVADNLKSSFTQPVSLSGGKLLLQKGVTLESTSFSQEA 455

Qy     480 GGVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNN 539
           | :: : :|  |          | :  |||: ::|:|: |:   | : |:           :
Db     456 GSLLGMDSGTTL---------STTAGSITITNLGINVDSL---GLKQPV----------S 493

Qy     540 YTADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFS 599
           || |:   :   ||:|||  ||  |||    |  |   : |: : : :     ::| |
Db     494 LTAKGASNKVIVSGKLNLIDIEGNI-YES----HMFSHDQLFSLLKITVDADVDTNVDIS 548

Qy     600 GL-NVP------HYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGY 652
           |  ||       ||:|| |   |         :: | |: ::|         ||   |:
Db     549 SLIPVPAEDPNSEYGFQGQWNVNW---------TTDTATNTKEAT-------ATWTKTGF 592

Qy     653 VPSPKHRSPLIANTLWGNMLLATESLKNSAEL--TPSDHP--FWGITGGGLGMMVYQDPR 708
           ||||: :| |: ||||| :     ||:   |:  |  :|   ||      :   :::
Db     593 VPSPERKSALVCNTLWG-VFTDIRSLQQLVEIGATGMEHKQGFW---VSSMTNFLHKTGD 648

Qy     709 ENHPGFHMRSSGYSAGMIAGQTHT-----FSLKFSQTYTKLNERY-AKNN---------- 752
           ||  ||   | ||   :| |  ||     |:  |   : :  : : | ||
Db     649 ENRKGFRHTSGGY---VIGGSAHTPKDDLFTFAFCHLFARDKDCFIAHNNSRTYGGTLFF 705

Qy     753 -----VSSKNYSCQGEMLFSLQ--EGF-----LLTKLVGLYSYGDHNCHHFYTQGENLTS 800
                :  :||   |   ||    | |     |   :   :|: |:     ||      |
Db     706 KHSHTLQPQNYLRLGRAKFSESAIEKFPREIPLALDVQVSFSHSDNRMETHYTSLPE--S 763

Qy     801 QGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALG---IYSSLSHFTEVGAYPRSFST 857
           :|:: :: : | :  |||     : | |   |:  :    :| | : | |  :  | ||
Db     764 EGSWSNECIAGGIGLDLPF-VLSNPHPLFKTFIPQMKVEMVYVSQNSFFESSSDGRGFSI 822

Qy     858 KTPLINVLVPIGVKGSFMNATQRPQAWTVELA--YQPVLYRQEPGIATQLLASKGIWFGS 915
             |:|: :|:| |  |:        ::| :|:  :  :|| |     |: |   |
Db     823 GR-LLNLSIPVGAK--FVQG-DIGDSYTYDLSGFFVSDVYRNNPQSTATLVMSPDSWKIR 878

Qy     916 GSPSSRHA 923
           |   || |
Db     879 GGNLSRQA 886


RESULT 13
PMPH_CHLTR
ID   PMPH_CHLTR     STANDARD;      PRT;  1016 AA.
AC   O84880;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable outer membrane protein pmpH precursor (Polymorphic membrane
DE   protein H).
GN   PMPH OR CT872.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D/UW-3/Cx;
RX   MEDLINE=99000809; PubMed=9784136;
RA   Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA   Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA   Davis R.W.;
RT   "Genome sequence of an obligate intracellular pathogen of humans:
RT   Chlamydia trachomatis.";
RL   Science 282:754-759(1998).
CC   -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC       (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE001360; AAC68470.1; -.
DR   PHCI-2DPAGE; O84880; -.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   InterPro; IPR003357; OMP.
DR   Pfam; PF02385; OMP; 1.
DR   Pfam; PF02415; DUF145; 1.
KW   Outer membrane; Signal; Multigene family; Complete proteome.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25   1016       PROBABLE OUTER MEMBRANE PROTEIN PMPH.
SQ   SEQUENCE   1016 AA;  107904 MW;  E691912C3A2BD6F7 CRC64;

  Query Match             11.7%;  Score 595;  DB 1;  Length 1016;
  Best Local Similarity   25.4%;  Pred. No. 9.4e-30;
```

```
  Matches  261;  Conservative  149;  Mismatches  404;  Indels  214;  Gaps   49;

Qy      42 NKISLTGDTHNLT-----NCYLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYF 96
           : :::||  | |:        | |  :| |         |  :|: |: |         : |
Db      69 DNLTITGQNHTLSFTDSQGPVLQNYAFISA--------GETLTLKDFSS--------LMF 112

Qy      97 AKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNT-------------------CCRP 137
           :||::    | |      :  || |    | ||| :|:                      |
Db     113 SKNVSCGEKGMI-----SGKTVSISGA-GEVIFWDNSVGYSPLSIVPASTPTPPAPAPAP 166

Qy     138 FTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQS 197
             ||: :  |:  |:|     :      : :: |  |   :                 |
Db     167 AASSSLSPTVSDARKGSIFSVET-----SLEISGVKKGVMF--------------DNNAG 207

Qy     198 CF--LFMDNICIQTNTAGKGGAIYAGTSNSFESNNCD---LFFINNACCAGGAIFSPICS 252
           |  :|  |     | || ||:  : |: ||   ||      |  | | | || ::
Db     208 NFGTVFRGN---SNNNAGSGGS-GSATTPSFTVKNCKGKVSFTDNVASCGGGVVYKGTVL 263

Qy     253 LTGNRGNIVFYNNRCFKNVETASSEASD--------GGAI-KVTTRLDVTGNRGRIFFSD 303
           | | | |  | : ::   :: : |         || |      :   ||:| | |
Db     264 FKDNEGGIFFRGNTAYDDLGILAATSRDQNTETGGGGGVICSPDDSVKFEGNKGSIVFDY 323

Qy     304 NITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIY---IDGTSNS-------KISAD 353
           |  |  ||:|     :|| :    | || |   ||||   || ::|          :|:
Db     324 NFAKGRGGSILTKEFSLVADDSVVFSNNTAEKGGGAIYAPTIDISTNGGSILFERNRAAE 383

Qy     354 RHAIIFNE-------NIVTNVTNA----NGTSTSANPPRRNAITVASSSGEILLGAGSSQ 402
             ||  :|       |:  : ::     :|  ||  |  |:|  : |    : | |
Db     384 GGAICVSEASSGSTGNLTLSASDGDIVFSGNMTSDRPGERSAARILSDGTTVSLNASGLS 443

Qy     403 NLIFYDPIEVSN--AGVS------------VSFNKEADQTGSVVFSGATVNSADFHQ-RN 447
           ||||||:  :|  || |                |: |:  :  |||:|:  ::  ::  |  |
Db     444 KLIFYDPVVQNNSAAGASTPSPSSSSMPGAVTINQSGN--GSVIFTAESLTPSEKLQVLN 501

Qy     448 LQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNSASNASI 507
           :  |  ||:| | | : : | ||    | | | |:||:|| ||     | |   | :
Db     502 STSNFPGALTVSGGELVVTEGATLTTGTITATSGRVTLGSGASLSAVAGAANN---NYTC 558

Qy     508 TLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYE 567
           |:  :|::| | |    :  :|  : |      | :: :|  |  |  |  : | | |
Db     559 TVSKLGIDLESFLTPNYKTAILGADGT-----VTVNSGSTLDL--VMESEAEVYDN-PLF 610

Qy     568 STDLTHALSSQPMLSISEAS-DNQLRSDDMDFSGLNVPHYGWQGLWTWGWAKTQDPEPAS 626
             ||      | :::| :|  | :  : :  :  :  |||:|| |:  | |   |   |
Db     611 VGSLT-----IPFVTLSSSSASNGVTKNSVTINDADAAHYGYQGSWSADWTK---PPLAP 662

Qy     627 SATITDPQKANRFHRTLLLTWLPAG----YVPSPKHRSPLIANTLW--GNMLLA-TESLK 679
           |    |   |    || ||| ||     |   |: :  |: |:||  |: |   |  ||
Db     663 DAKGMVPPNTN---NTLYLTWRPASNYGEYRLDPQRKGELVPNSLWVAGSALRTFTNGLK 719

Qy     680 -----NSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTH--- 731
                 | |          ||  :    :::  ||  |  |: |      :|
Db     720 EHYVSRDVGFVASLH--------ALGDYILNYTQDDRDGFLARYGGFQA---TAASHYEN 768

Qy     732 --TFSLKFSQTY--TKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGLY--SYGD 785
             | : | |  |   ||    |:|:  :    || |     ::     |: |  :  :||
Db     769 GSIFGVAFGQLYGQTKSRMYYSKDAGNMTMLSCFGRSYVDIKG----TETVMYWETAYG- 823

Qy     786 HNCHHFYTQGENLTSQGTFRSQ---------TMGGA----VFFDLPMKPFGSTHILTA-- 830
           :: |  :||  |  :|    |:          ||    : : :| : |   : ||
Db     824 YSVHRMHTQYFNDKTQKFDHSKCHWHNNNYYAFVGAEHNFLEYCIPTRQFARDYELTGFM 883

Qy     831 PFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQR--PQAWTVEL 888
           |  | |  ||     | |:  | |:  :   |: :|||:    ::  :|  |   |: :
Db     884 RFEMAGGWSSST---RETGSLTRYFARGSG-HNMSLPIGIVAHAVSHVRRSPPSKLTLNM 939

Qy     889 AYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFY 948
           |:| ::|  |    :::|:       ||| :|||   ::        ||:  : |
Db     940 GYRPDIWRVTPHCNMEIIANGVKTPIQGSPLARHAFFLEVHD--------TLYIHHFG-- 989

Qy     949 SSSTFCNY 956
             : ||
Db     990 --RAYMNY 995
```

```
RESULT 14
PM10_CHLPN
ID   PM10_CHLPN     STANDARD;      PRT;   928 AA.
AC   Q9RB65; Q9RB64; Q9S6P2; O86163;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable outer membrane protein pmp10 precursor (Polymorphic membrane
DE   protein 10) (Outer membrane protein 5).
GN   PMP10 OR OMP5 OR CP0303.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VR1310;
RX   MEDLINE=20007584; PubMed=10539856;
RA   Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
RA   Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT   "Molecular biology of Chlamydia pneumoniae surface proteins and their
RT   role in immunopathogenicity.";
RL   Am. Heart J. 138:S491-S495(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AR39;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
RN   [4]
RP   SEQUENCE OF 1-914 FROM N.A.
RC   STRAIN=CWL029/VR-1310;
RX   MEDLINE=99081766; PubMed=9864239;
RA   Knudsen K., Madsen A.S., Mygind P., Christiansen G., Birkelund S.;
RT   "Identification of two novel genes encoding 97- to 99-kilodalton outer
RT   membrane proteins of Chlamydia pneumoniae.";
RL   Infect. Immun. 67:375-383(1999).
CC   -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC       (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ133034; CAB37071.1; -.
DR   EMBL; AE002192; AAF38160.1; -.
DR   EMBL; AP002546; BAA98657.1; -.
DR   EMBL; AJ001311; CAA04671.1; -.
DR   PHCI-2DPAGE; O86163; -.
DR   TIGR; CP0303; -.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   InterPro; IPR003357; OMP.
```

```
DR   Pfam; PF02385; OMP; 1.
DR   Pfam; PF02415; DUF145; 2.
KW   Outer membrane; Signal; Multigene family.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26    928       PROBABLE OUTER MEMBRANE PROTEIN PMP10.
SQ   SEQUENCE   928 AA;  97229 MW;  0590D5206A1DD0E1 CRC64;

  Query Match             11.5%;  Score 586.5;  DB 1;  Length 928;
  Best Local Similarity   25.6%;  Pred. No. 2.8e-29;
  Matches  264;  Conservative  143;  Mismatches  377;  Indels  249;  Gaps   48;

Qy      1 MKKAFFFFLIGNSL----------SGLAREV-PSRIF--------LMPNSVPDPTKESLS 41
          ||  | : :: ::|          :  |  : ||  |           | :       :|:
Db      1 MKSQFSWLVLSSTLACFTSCSTVFAATAENIGPSDSFDGSTNTGTYTPKNTTTGIDYTLT 60

Qy     42 NKISL--TGDTHNLT-NCYLDNLR--------YILAILQ-KTPNEGAAVTIT-------- 81
            |:|   ||:  ||  |: |           | |: |  |:  ||||:::|
Db     61 GDITLQNLGDSAALTKGCFSDTTESLSFAGKGYSLSFLNIKSSAEGAALSVTTDKNLSLT 120

Qy     82 --DYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFT 139
            |:|       |      ||   ||:      :      |  | ::|: : |
Db    121 GFSSLTFLAAPSSVI-----TTPSGKGAVKCGGDLT-----FDNNGTILFKQDYC----- 165

Qy    140 SSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSCF 199
                    ||||  :|| : ::   : |  | |
Db    166 ----------EENGGAISTKNLSLKNSTGSISFEGNKS---------------------- 193

Qy    200 LFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINN-ACCAGGAIFSP-ICSLTGNR 257
                  : |  |||||  |  :    :|     | || |  |||||  |   |::|||
Db    194 ---------SATGKKGGAICATGTVDITNNTAPTLFSNNIAEAAGGAINSTGNCTITGNT 244

Qy    258 GNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYAPV 317
           ::||  |       : :: | :|||:     : ::||:  : || |     ||||||
Db    245 -SLVFSEN-------SVTATAGNGGALSGDADVTISGNQS-VTFSGNQAVANGGAIYAKK 295

Qy    318 VTLVD--NGPTYFINNIAN----NKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNAN 371
           :||     |   | |||        |||| |       :||:    | || | :
Db    296 LTLASGGGGGISFSNNIVQGTTAGNGGAISILAAGECSLSAEAGDITFNGNAI------- 348

Qy    372 GTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVS---VSFNK----- 423
             :|:    :||:| : |::    | | |   :: |||||  : |  |    :: ||
Db    349 -VATTPQTTKRNSIDIGSTAKITNLRAISGHSIFFYDPITANTAADSTDTLNLNKADAGN 407

Qy    424 EADQTGSVVFSGATVNS-----ADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQ 478
            | :||:|||| ::      ||    || :    |:||: | | ::    |    |||
Db    408 STDYSGSIVFSGEKLSEDEAKVAD----NLTSTLKQPVTLTAGNLVLKRGVTLDTKGFTQ 463

Qy    479 T-GGVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNS 537
          | |  | :  |  |          ||   :|  : : : | |  | :: :
Db    464 TAGSSVIMDAGTTL---------KASTEEVTLTGLSIPVDS-LGEGKKVVI--------- 504

Qy    538 NNYTADTAATFSLSDVKLS----LIDDYGNSPYESTDL--THALSSQPMLSISEASDNQL 591
                ||: :  :| ||    |:|: ||: ||: ||  |    |   : ::  |:   :
Db    505 -------AASAASKNVALSGPILLLDNQGNA-YENHDLGKTQDFSFVQLSALGTATTTDV 556

Qy    592 RSDDMDFSGLNVP------HYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLL 645
           :         ||      |||:|| |   |       | :               :|  |
Db    557 PA---------VPTVATPTHYGYQGTWGMTWVDDTASTPKT-------------KTATL 593

Qy    646 TWLPAGYVPSPKHRSPLIANTLWGNM--LLATES-LKNSAELTPSDHPFWGITGGGLGMM 702
            |   ||:|:|: : ||: |:|||:    : | :  :: ||    ||  ||     |:
Db    594 AWTNTGYLPNPERQGPLVPNSLWGSFSDIQAIQGVIERSALTLCSDRGFW---AAGVANF 650

Qy    703 VYQDPRENHPGFHMRSSGYSAGMIAGQT---HTFSLKFSQTYTKLNERYAKNNVSSKNYS 759
          : :| :     :  :| ||: |  | ||    :  |  | | :        |: : :||::
Db    651 LDKDKKGEKRKYRHKSGGYAIGG-AAQTCSENLISFAFCQLFGS-----DKDFLVAKNHT 704

Qy    760 CQGEMLFSLQE--------GFLLTKLVG-------------LYSYGDHNCHHFYT-QGEN 797
                | :|         | || || |               ||:  ::    ||   |
Db    705 DTYAGAFYIQHITECSGFIGCLLDKLPGSWSHKPLVLEGQLAYSHVSNDLKTKYTAYPEV 764

Qy    798 LTSQGTFRSQTMGGAVFFDLP--MKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSF 855
            | |      | ||     |  :  | :     ||::     |     |:| |   |||
```

```
Db    765 KGSWGNNAFNMMLGASSHSYPEYLHCFDT----YAPYIKLNLTYIRQDSFSEKGTEGRSF 820

Qy    856 STKTPLINVLVPIGVK-GSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFG 914
                | |: :|||||   | :      :: : |:| | | | :|   | |: |   |
Db    821 DDSN-LFNLSLPIGVKFEKFSDCND--FSYDLTLSYVPDLIRNDPKCTTALVISGASWET 877

Qy    915 SGSPSSRHAMSYK 927
                :  :| |:  :
Db    878 YANNLARQALQVR 890


RESULT 15
PM19_CHLPN
ID   PM19_CHLPN     STANDARD;      PRT;   947 AA.
AC   Q9Z813; Q9JSE2;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Probable outer membrane protein pmp19 precursor (Polymorphic membrane
DE   protein 19).
GN   PMP19 OR CPN0539 OR CP0213.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CWL029;
RX   MEDLINE=99206606; PubMed=10192388;
RA   Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA   Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT   "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL   Nat. Genet. 21:385-389(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AR39;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20298986; PubMed=10839753;
RA   Shirai M., Hirakawa H., Ouchi K., Tabuchi M., Kishi F., Kimoto M.,
RA   Takeuchi H., Nishida J., Shibata K., Fujinaga R., Yoneda H.,
RA   Matsushima H., Tanaka C., Furukawa S., Miura K., Nakazawa A.,
RA   Ishii K., Shiba T., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of outer membrane protein genes omp and pmp in the whole
RT   genome sequences of Chlamydia pneumoniae isolates from Japan and the
RT   United States.";
RL   J. Infect. Dis. 181 Suppl 3:S524-S527(2000).
CC   -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC       (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE001638; AAD18679.1; -.
DR   EMBL; AE002181; AAF38083.1; -.
DR   EMBL; AP002547; BAA98745.1; -.
DR   HSSP; Q90121; 1KPT.
DR   TIGR; CP0213; -.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   InterPro; IPR003357; OMP.
DR   Pfam; PF02385; OMP; 1.
DR   Pfam; PF02415; DUF145; 1.
KW   Outer membrane; Signal; Multigene family; Complete proteome.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    947       PROBABLE OUTER MEMBRANE PROTEIN PMP19.
FT   CONFLICT    453    453       E -> D (IN REF. 3).
SQ   SEQUENCE   947 AA;  103642 MW;  20CE1DEEE1606DFF CRC64;

  Query Match             11.5%;  Score 586;  DB 1;  Length 947;
  Best Local Similarity   25.2%;  Pred. No. 3.1e-29;
  Matches  250;  Conservative  173;  Mismatches  390;  Indels  178;  Gaps   47;

Qy     14 LSGLAREVPSRIFLMPNSVPDPTKESLSNKI---SLTGDTHNLTNCYLDNLRYILAILQK 70
          |||:       :| : :| |  |  ||          |:: :       ||:  |  :
Db     30 LSGIHSGEDLELFTLRSSSPTKTTYSLRKDFIVCDFAGNSIHKPGAAFLNLKGDLFFINS 89

Qy     71 TPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVE-------IRDT 123
          ||   ||:|  :       :  |::   |:|        :   :|  :|       :  |
Db     90 TPL--AALTFKNI--HLGARGAGLFSESNVT--------FKGLHSLVLENNESWGGVLTT 137

Qy    124 IGPVIFENNT--CCRPFTSSNPNAAVNKIREGGAIHAQN-----LYINHNHDVVGFMKNF 176
           | : | |||   |:   |  |         |||:  |      |:   |   : |:||
Db    138 SGDLSFINNTSVLCQNNISYGP---------GGALLLQGRKSKALFFRDNRGTILFLKNK 188

Qy    177 SYVR-------GGAISTANTFVVSENQSCFLFMDN--ICIQTNTAGKGGAIYAGT-SNSF 226
          :  :       |||:|: :        |   | ||  |  | |    |||||    : :|
Db    189 AVNQDESHPGYGGAVSSISP------GSPITFADNQEILFQENEGELGGAIYNDQGAITF 242

Qy    227 ESNNCDLFFINNACCAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKV 286
          |:|     | :|    |||::|  |:|    |: :|          | :: |  ||||
Db    243 ENNFQTTSFFSNKASFGGAVYSRYCNLYSQWGDTLF----------TKNAAAKVGGAIH- 291

Qy    287 TTRLDVTGNRGRIFFSDNITKNYGGAIYAPVVTLVD-NGPTYFINNIA-NNKGGAIYIDG 344
          : :   :| | | :| :   ||||    |  ::  ||  |||| |    |||||:
Db    292 ADYVHIRDCKGSIVFEEN-SATAGGAIAVNAVCDINAQGPVRFINNSALGLNGGAIYMQA 350

Qy    345 T-SNSKISADRHAIIFNENIVTNV--TNANGTSTSANPPRRNAITVASSSGEILLGAGSS 401
          | |  :: |::  | |  | | :   :: | ||   |    ||||:  :  |  | |
Db    351 TGSILRLHANQGDIEFCGNKVRSQFHSHINSTSNFTN----NAITIQGAPREFSLSANEG 406

Qy    402 QNLIFYDPI----EVSNAGVSVSFNKEADQTGSVVFSGATVNSADFHQRNLQTKTP---A 454
            : |||||    |  |: : ::  :  :  |:|:|||| ::    |::  : ||
Db    407 HRICFYDPIISATENYNS-LYINHQRLLEAGGAVIFSGARLSPE--HKKENKNKTSIINQ 463

Qy    455 PLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNSASNASITLKHIGL 514
          |: | :| | ||  | | |  | | ||:::|| |: |:      | ::    | : ::|
Db    464 PVRLCSGVLSIEGGAILAVRSFYQEGGLLALGPGSKLTT----QGKNSEKDKIVITNLGF 519

Qy    515 NLSSILKSG-AEIPLLWVEPTNNSNNYTADTAATFSLSDVKL------SLID--DYGNSP 565
          || ::  |  |||               |   |:  :| |        |  :  :| : |
Db    520 NLENLDSSDPAEI-------------RATEKASIEISGVPRVYGHTESFYENHEYASKP 565

Qy    566 YESTDLTHALSSQPMLSISEASDNQLRS---DDMDFSGLNVPHYGWQGLWTWGWAKTQDP 622
          | :: :   ||:: :::     :  :::    : :: |     ||:|| | : |:
Db    566 YTTSII---LSAKKLVTAPSRPEKDIQNLIIAESEYMG-----YGYQGSWEFSWSPNDTK 617

Qy    623 EPASSATITDPQKANRFHRTLLLTWLPAG-YVPSPKHRSPLIANTLWGNML-------LA 674
          |              :|:: :| | | :   || |   |  |||             :
Db    618 E---------------KKTIIASWTPTGEFSLDPKRRGSFIPTTLWSTFSGLNIASNIV 661

Qy    675 TESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFS 734
            :  |::|: |  |    : || :  :: |:|:::     :: :|:: |  |    :|:

Db    662 NNNYLNNSEVIPLQH--LCVFGGPVYQIMEQNPKQSSNNLLVQHAGHNVG--ARIPFSFN 717

Qy    735 LKFSQTYTKLNERYAKNNVSSKNYSCQGEMLF---SLQEGFLLTKLVGLYSYGDHN---C 788
             |   |:|    :: ||: |:::   ::|    || : :    |   :|| : :
Db    718 TILSAALTQLFSSSSQQNVADKSHA---QILIGTVSLNKSWQALSLRSSFSYTEDSQVMK 774

Qy    789 HHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSL--SHFT 846
          | |  :|   ||:|::|:    |:|       | |  ::   ||:     |: |  : |
Db    775 HVFPYKG---TSRGSWRNYGWSGSVGMSYAY-PKGIRYLKMTPFVDLQ--YTKLVQNPFV 828

Qy    847 EVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAW-TVELAYQPVLYRQEPGIATQL 905
          | |  || ||: : : |: :|||:         |   :  |  :|   | |  |  :  |
Db    829 ETGYDPRYFSS-SEMTNLSLPIGIALEMRFIGSRSSLFLQVSTSYIKDLRRVNPQSSASL 887

Qy    906 LASKGIWFGSGSPSSRHAMS--------YKI 928
          : :   |   | |  : |::        |||
Db    888 VLNHYTWDIQGVPLGKEALNITLNSTIKYKI 918


Search completed: February  4, 2003, 20:55:58
Job time : 44 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2003, 20:55:07 ; Search time 53 Seconds
(without alignments)
1750.372 Million cell updates/sec

Title: US-09-677-752-2
Perfect score: 5086
Sequence: 1 MKKAFFFFLIGNSLSGLARE.....GFYSSSTFCNYLNGEIALRF 965

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4989.5	98.1	964	2	E71460	probable outer mem
2	3648.5	71.7	976	2	F81722	polymorphic membra
3	1120.5	22.0	938	2	F86548	polymorphic outer
4	1120.5	22.0	938	2	H72074	polymorphic membra
5	1060.5	20.9	952	2	D81593	polymorphic membra
6	1052	20.7	934	2	A72075	polymorphic outer
7	1046	20.6	934	2	G86548	polymorphic outer
8	951.5	18.7	946	2	C86549	polymorphic outer
9	951.5	18.7	946	2	D81594	polymorphic membra
10	948.5	18.6	946	2	C72075	polymorphic outer
11	866.5	17.0	1034	2	F71460	probable outer mem
12	793	15.6	1025	2	G81722	polymorphic membra
13	655.5	12.9	1013	2	G71460	probable outer mem
14	645	12.7	936	2	C72078	polymorphic outer
15	644	12.7	936	2	C86546	polymorphic outer
16	644	12.7	936	2	B81591	polymorphic membra
17	620.5	12.2	987	2	H81722	polymorphic membra
18	620	12.2	930	2	D86546	polymorphic outer
19	620	12.2	930	2	A81591	polymorphic membra
20	617	12.1	930	2	D72078	polymorphic outer
21	597	11.7	928	2	H86546	polymorphic outer
22	597	11.7	928	2	D72077	polymorphic outer
23	597	11.7	949	2	F81591	polymorphic membra
24	595	11.7	1016	2	H71460	probable outer mem
25	587	11.5	947	2	G86557	polymorphic membra
26	586.5	11.5	928	2	G86546	polymorphic outer
27	586.5	11.5	928	2	G81591	polymorphic membra
28	586	11.5	947	2	D72067	polymorphic membra
29	585.5	11.5	928	2	E86546	polymorphic outer
30	585.5	11.5	928	2	B72077	polymorphic membra
31	583	11.5	978	2	G72076	polymorphic outer
32	583	11.5	978	2	B81593	polymorphic membra
33	583	11.5	978	2	C86547	polymorphic outer
34	575	11.3	973	2	B86547	polymorphic outer
35	575	11.3	973	2	F72076	polymorphic outer
36	575	11.3	995	2	C81593	polymorphic membra
37	570.5	11.2	922	2	B72131	polymorphic outer
38	570.5	11.2	922	2	E86491	polymorphic outer
39	569.5	11.2	922	2	F81539	polymorphic membra
40	567	11.1	1276	2	B86546	polymorphic outer
41	567	11.1	1276	2	C81591	polymorphic membra
42	564	11.1	983	2	A81723	polymorphic membra
43	562	11.0	1407	2	B72078	polymorphic outer
44	546.5	10.7	986	2	B81675	polymorphic membra
45	530.5	10.4	772	2	H86492	Pmp_3 [imported] -

ALIGNMENTS

RESULT 1
E71460
probable outer membrane protein E - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C;Species: Chlamydia trachomatis
C;Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
C;Accession: E71460
R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche
Science 282, 754-759, 1998
A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t
A;Reference number: A71570; MUID:99000809; PMID:9784136
A;Accession: E71460
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-964 <ARN>
A;Cross-references: GB:AE001360; GB:AE001273; NID:g3329342; PIDN:AAC68467.1; PID:g332
A;Experimental source: serotype D, strain UW-3/Cx
C;Genetics:
A;Gene: pmpE

Query Match 98.1%; Score 4989.5; DB 2; Length 964;
Best Local Similarity 98.3%; Pred. No. 3.8e-298;
Matches 949; Conservative 7; Mismatches 8; Indels 1; Gaps 1;

```
Qy       1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 60

Qy      61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120

Qy     121 RDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVR 180
           |||||||||||||||  ||  || ||  :|||||||||||||||||||||||||||||||:
Db     121 RDTIGPVIFENNTCCRLFTWRNPYAA-DKIREGGAIHAQNLYINHNHDVVGFMKNFSYVQ 179

Qy     181 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 239

Qy     241 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     240 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 299

Qy     301 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFN 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     300 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFN 359

Qy     361 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     360 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 419

Qy     421 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 480
```

```
Db    420 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 479

Qy    481 GVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNY 540

Db    480 GVVSLGNGAVLSCYKNGTGDSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNY 539

Qy    541 TADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSG 600

Db    540 TADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLQSENIDFSG 599

Qy    601 LNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRS 660

Db    600 LNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRS 659

Qy    661 PLIANTLWGNMLLATESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSG 720

Db    660 PLIANTLWGNMLLATESLKNSAELTPSGHPFWGITGGGLGMMVYQDPRENHPGFHMRSSG 719

Qy    721 YSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGL 780

Db    720 YSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSLQEGFLLTKLVGL 779

Qy    781 YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYS 840

Db    780 YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYS 839

Qy    841 SLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPG 900

Db    840 SLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATHRPQAWTVELAYQPVLYRQEPG 899

Qy    901 IATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGE 960

Db    900 IAAQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGE 959

Qy    961 IALRF 965

Db    960 IALRF 964


RESULT 2
F81722
polymorphic membrane protein E/F family TC0261 [imported] - Chlamydia muridarum (strain
C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
C;Accession: F81722
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: F81722
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-976 <TET>
A;Cross-references: GB:AE002293; GB:AE002160; NID:g7190298; PIDN:AAF39130.1; PID:g719030
A;Experimental source: strain Nigg (MoPn)
C;Genetics:
A;Gene: TC0261

  Query Match             71.7%;  Score 3648.5;  DB 2;  Length 976;
  Best Local Similarity   70.5%;  Pred. No. 7.3e-216;
  Matches  689;  Conservative  121;  Mismatches  154;  Indels   13;  Gaps    3;

Qy      1 MKKAFFFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLTNCYLDN 60

Db      1 MKKLFFFVLIGSSILGFTREVPPSILLKPILNPYHMTGLFFPKVNLLGDTHNLTDYHLDN 60

Qy     61 LRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEI 120

Db     61 LKCILACLQRTPYEGAAFTVTDYLGFSDTQKDGIFCFKNLTPESGGVIGSPTQNTPTIKI 120

Qy    121 RDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVR 180

Db    121 HNTIGPVLFENNTCHRLWTQTDPENEGNKAREGGAIHAGDVYISNNQNLVGFIKNFAYVQ 180

Qy    181 GGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFINNAC 240

Db    181 GGAIS-ANTFAYKENKSSFLCLNNSCIQTKTGGKGGAIYVSTSCSFENNNKDLLFIQNSG 239

Qy    241 CAGGAIFSPICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 300

Db    240 CAGGAIFSPTCSLIGNQGDIVFYSNHGFKNVDNATNESGDGGAIKVTTRLDITNNGSQIF 299

Qy    301 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFN 360

Db    300 FSDNISRNFGGAIHAPCLHLVGNGPTYFTNNIANHTGGAIYITGTETSKISADHHAIIFD 359

Qy    361 ENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEVSNAGVSVS 420

Db    360 NNISANATNADGSSSNTNPPHRNAITMDNSAGGIELGAGKSQNLIFYDPIQVTNAGVTVD 419

Qy    421 FNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTG 480

Db    420 FNKDASQTGCVVFSGATVLSADISQANLQTKTPATLTLSHGLLCIEDRAQLTVNNFTQTG 479

Qy    481 GVVSLGNGAVLSCYKNGAGN--------SASNASITLKHIGLNLSSILKSGAEIPLLWVE 532

Db    480 GIVALGNGAVLSSYQHSTTDATQTPPTTTTTDASVTLNHIGLNLPSILKDGAEMPLLWVE 539

Qy    533 PTN----NSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASD 588

Db    540 PISTTQGNTTTYTSDTAASFSLNGATLSLIDEDGNSPYENTDLSRALYAQPMLAISEASD 599

Qy    589 NQLRSDDMDFSGLNVPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWL 648

Db    600 NQLQSESMDFSKVNVPHYGWQGLWTWGWAKTENPTTTPPATITDPKKANQFHRTLLLTWL 659

Qy    649 PAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSDHPFWGITGGGLGMMVYQDPR 708

Db    660 PAGYIPSPKHKSPLIANTLWGNILFATENLKNSSGQELLDRPFWGITGGGLGMMVYQEPR 719

Qy    709 ENHPGFHMRSSGYSAGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFSL 768

Db    720 KDHPGFHMHTSGYSAGMITGNTHTFSLRFSQSYTKLNERYAKNYVSSKNYSCQGEMLLSL 779

Qy    769 QEGFLLTKLVGLYSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHIL 828

Db    780 QEGLMLTKLIGLYSYGNHNSHHFYTQGEDLSSQGEFHSQTFGGAVFFDLPLKPFGRTHIL 839

Qy    829 TAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVEL 888

Db    840 TAPFLGAIGMYSKLSSFTEVGAYPRTFITETPLINVLIPIGVKGSFMNATHRPQAWTVEL 899

Qy    889 AYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFY 948

Db    900 AYQPVLYRQEPSISTQLLAGKGMWFGHGSPASRHALAYKISQKTQLLRFATLQLQYHGYY 959

Qy    949 SSSTFCNYLNGEIALRF 965

Db    960 SSSTFCNYLNGEVSLRF 976


RESULT 3
F86548
polymorphic outer membrane protein E family [imported] - Chlamydophila pneumoniae (st
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C;Accession: F86548
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;
Nucleic Acids Res. 28, 2311-2314, 2000
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;Reference number: A86491; MUID:20330349; PMID:10871362
A;Accession: F86548
A;Status: preliminary
A;Molecule type: DNA
```

```
A;Residues: 1-938 <STO>
A;Cross-references: GB:BA000008; NID:g8978836; PIDN:BAA98772.1; GSPDB:GN00142
A;Experimental source: strain J138
C;Genetics: -
A;Gene: pmp_15

  Query Match             22.0%;  Score 1120.5;  DB 2;  Length 938;
  Best Local Similarity   30.9%;  Pred. No. 8.9e-61;
  Matches  308;  Conservative  163;  Mismatches  418;  Indels  109;  Gaps   31;

Qy      6 FFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLT---------NC 56
          | |:: |   ||   : : | | |     |             | |||           |
Db     12 FTFVLAN--EGLQLPLETYITLSPEYQAAPQ----------VGFTHNQNQDLAIVGNHND 59

Qy     57 YLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSP 116
          :: : :|        : | |:|  : |         ::| ||: | ||||| ||: |
Db     60 FILDYKYY-------RSNGGALTCKNLL--ISENIGNVFFEKNVCPNSGGAI-YAAQNC- 108

Qy    117 TVEIRDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNF 176
              ||     :|       | || |    :  |||: | |  | :|     |: |
Db    109 ------TISK--NQNYAFTTNLVSDNPTATAGSLL-GGALFAINCSITNNLGQGTFVDNL 159

Qy    177 SYVRGGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFI 236
          :  :|||: |     : :|:   :   |  :  |:   || ||:| | : | |:  :
Db    160 ALNKGGALYTETNLSIKDNKGPIIIKQNRAL--NSDSLGGGIYSGNSLNIEGNSGAIQIT 217

Qy    237 NNACCAGGAIFS-PICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGN 295
          :|:  :|| |||    ::: |:  |    |  | |     |  : ||    ||   :  |
Db    218 SNSSGSGGGIFSTQTLTISSNKKLIEISENSAFAN--NYGSNFNPGGGGLTTTFCTILNN 275

Qy    296 RGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYID-GTSNSK--ISA 352
          |  : |::| ::: ||||:|  : : :||| ||:|| |    |  : :  |: |    :||
Db    276 REGVLFNNNQSQSNGGAIHAKSIIIKENGPVYFLNNTATRGGALLNLSAGSGNGSFILSA 335

Qy    353 DRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEV 412
          |   |||| |          :  : ||| ||||  :: :  : :||     ::||||||
Db    336 DNGDIIFNNN--------TASKHALNPPYRNAIH-STPNMNLQIGARPGYRVLFYDPIEH 386

Qy    413 S-NAGVSVSFNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQL 471
          :   : || |   ||:|:|||  |:     : | :       |  | | :|| | |
Db    387 ELPSSFPILFNFETGHTGTVLFSGEHVHQNFTDEMNFFSYLRNTSELRQGVLAVEDGAGL 446

Qy    472 TVNRFTQTGGVVSLGNGAVLSCYKNGAGN--------SASNASITLKHIGLNLSSILKSG 523
          :| | || : || |||::     ||          :   ::||| || ::| |||
Db    447 ACYKFFQRGGTLLLGQGAVITT----AGTIPTPSSTPTTVGSTITLNHIAIDLPSILSFQ 502

Qy    524 AEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSI 583
          |: | :|: ||   : || |:  | ::|   |:| :     ||:| ||:|:|   |:| |
Db    503 AQAPKIWIYPTKTGSTYTEDSNPTITISGT-LTLRNSNNEDPYDSLDLSHSLEKVPLLYI 561

Qy    584 SEASDNQLRSDDMDFSGLNV-PHYGWQGLWTWGWAKTQDPEPASSATITDPQK---ANRF 639
          : :  :: |  :| | ||   |||:||:|:  | :|         |||:|     ||
Db    562 VDVAAQKINSSQLDLSTLNSGEHYGYQGIWSTYWVET--------TTITNPTSLLGANTK 613

Qy    640 HRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSDHPFWGITGG-- 697
          |: |   | | || | |: |   | | || :   |   | | : :          |    |
Db    614 HKLLYANWSPLGYRPHPERRGEFITNALWQSAYTALAGLHSLSS--------WDEEKGHA 665

Qy    698 ----GLGMMVYQDPRENHPGFHMRSSGYSAGMIA--GQTHTFSLKFSQTYTKLNERYAKN 751
              |:|::|:|  :    ||    :|||||   |   |:  ||| |:| ::|  |  ::|
Db    666 ASLQGIGLLVHQKDKNGFKGFRSHMTGYSATTEATSSQSPNFSLGFAQFFSKAKEHESQN 725

Qy    752 NVSSKNY---SCQGEMLFSLQEGFLLTKLVGL-YSYGDHNCHHFYTQGENLTSQGTFRSQ 807
          : || :|    |   ||  :|  |:  | | | :   : |  |        |||:| :
Db    726 STSSHHYFSGMCIENTLF--KEWIRLS--VSLAYMFTSEHTHTMYQGLLEGNSQGSFHNH 781

Qy    808 TMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVP 867
          |: ||:      :| | : :   ||: || |  :|: | | | : | ||   || :| :|
Db    782 TLAGALSCVFLPQPHGES-LQIYPFITALAIRGNLAAFQESGDHAREFSLHRPLTDVSLP 840

Qy    868 IGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYK 927
          :|:: |: |  : |  |  |:::|:  ||||:| : ::|| |:| |    :| : :|:  |
Db    841 VGIRASWKNHHRVPLVWLTEISYRSTLYRQDPELHSKLLISQGTWTTQATPVTYNALGIK 900

Qy    928 ISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
          :    |    :||   |    ||||  :|||    :||
Db    901 VKNTMQVFPKVTLSLDYSADISSSTLSHYLNVASRMRF 938


RESULT 4
H72074
polymorphic membrane protein E/F family CP0286 [imported] - Chlamydophila pneumoniae
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C;Accession: H72074; E81593
R;Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,
Nature Genet. 21, 385-389, 1999
A;Title: Comparative genomes of Clamydia pneumoniae and C. trachomatis.
A;Reference number: A72000; MUID:99206606; PMID:10192388
A;Accession: H72074
A;Molecule type: DNA
A;Residues: 1-938 <ARN>
A;Cross-references: GB:AE001631; GB:AE001363; NID:g4376750; PIDN:AAD18608.1; PID:g437
A;Experimental source: strain CWL029
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: E81593
A;Molecule type: DNA
A;Residues: 1-938 <REA>
A;Cross-references: GB:AE002190; GB:AE002161; NID:g7189209; PIDN:AAF38143.1; PID:g718
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: pmp_15; CP0286

  Query Match             22.0%;  Score 1120.5;  DB 2;  Length 938;
  Best Local Similarity   30.9%;  Pred. No. 8.9e-61;
  Matches  308;  Conservative  163;  Mismatches  418;  Indels  109;  Gaps   31;

Qy      6 FFFLIGNSLSGLAREVPSRIFLMPNSVPDPTKESLSNKISLTGDTHNLT---------NC 56
          | |:: |   ||   : : | | |     |             | |||           |
Db     12 FTFVLAN--EGLQLPLETYITLSPEYQAAPQ----------VGFTHNQNQDLAIVGNHND 59

Qy     57 YLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSP 116
          :: : :|        : | |:|  : |         ::| ||: | ||||| ||: |
Db     60 FILDYKYY-------RSNGGALTCKNLL--ISENIGNVFFEKNVCPNSGGAI-YAAQNC- 108

Qy    117 TVEIRDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNF 176
              ||     :|       | || |    :  |||: | |  | :|     |: |
Db    109 ------TISK--NQNYAFTTNLVSDNPTATAGSLL-GGALFAINCSITNNLGQGTFVDNL 159

Qy    177 SYVRGGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSNSFESNNCDLFFI 236
          :  :|||: |     : :|:   :   |  :  |:   || ||:| | : | |:  :
Db    160 ALNKGGALYTETNLSIKDNKGPIIIKQNRAL--NSDSLGGGIYSGNSLNIEGNSGAIQIT 217

Qy    237 NNACCAGGAIFS-PICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGN 295
          :|:  :|| |||    ::: |:  |    |  | |     |  : ||    ||   :  |
Db    218 SNSSGSGGGIFSTQTLTISSNKKLIEISENSAFAN--NYGSNFNPGGGGLTTTFCTILNN 275

Qy    296 RGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYID-GTSNSK--ISA 352
          |  : |::| ::: ||||:|  : : :||| ||:|| |    |  : :  |: |    :||
Db    276 REGVLFNNNQSQSNGGAIHAKSIIIKENGPVYFLNNTATRGGALLNLSAGSGNGSFILSA 335

Qy    353 DRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIEV 412
          |   |||| |          :  : ||| ||||  :: :  : :||     ::||||||
Db    336 DNGDIIFNNN--------TASKHALNPPYRNAIH-STPNMNLQIGARPGYRVLFYDPIEH 386

Qy    413 S-NAGVSVSFNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGFLCIEDHAQL 471
          :   : || |   ||:|:|||  |:     : | :       |  | | :|| | |
Db    387 ELPSSFPILFNFETGHTGTVLFSGEHVHQNFTDEMNFFSYLRNTSELRQGVLAVEDGAGL 446

Qy    472 TVNRFTQTGGVVSLGNGAVLSCYKNGAGN--------SASNASITLKHIGLNLSSILKSG 523
          :| | || : || |||::     ||          :   ::||| || ::| |||
Db    447 ACYKFFQRGGTLLLGQGAVITT----AGTIPTPSSTPTTVGSTITLNHIAIDLPSILSFQ 502
```

```
Qy    524 AEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSI 583
Db    503 AQAPKIWIYPTKTGSTYTEDSNPTITISGT-LTLRNSNNEDPYDSLDLSHSLEKVPLLYI 561

Qy    584 SEASDNQLRSDDMDFSGLNV-PHYGWQGLWTWGWAKTQDPEPASSATITDPQK---ANRF 639
Db    562 VDVAAQKINSSQLDLSTLNSGEHYGYQGIWSTYWVET--------TTITNPTSLLGANTK 613

Qy    640 HRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSDHPFWGITGG-- 697
Db    614 HKLLYANWSPLGYRPHPERRGEFITNALWQSAYTALAGLHSLSS--------WDEEKGHA 665

Qy    698 ----GLGMMVYQDPRENHPGFHMRSSGYSAGMIA--GQTHTFSLKFSQTYTKLNERYAKN 751
Db    666 ASLQGIGLLVHQKDKNGFKGFRSHMTGYSATTEATSSQSPNFSLGFAQFFSKAKEHESQN 725

Qy    752 NVSSKNY---SCQGEMLFSLQEGFLLTKLVGL-YSYGDHNCHHFYTQGENLTSQGTFRSQ 807
Db    726 STSSHHYFSGMCIENTLF--KEWIRLS--VSLAYMFTSEHTHTMYQGLLEGNSQGSFHNH 781

Qy    808 TMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVP 867
Db    782 TLAGALSCVFLPQPHGES-LQIYPFITALAIRGNLAAFQESGDHAREFSLHRPLTDVSLP 840

Qy    868 IGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYK 927
Db    841 VGIRASWKNHHRVPLVWLTEISYRSTLYRQDPELHSKLLISQGTWTTQATPVTYNALGIK 900

Qy    928 ISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
Db    901 VKNTMQVFPKVTLSLDYSADISSSTLSHYLNVASRMRF 938
```

RESULT 5
D81593
polymorphic membrane protein E/F family CP0285 [imported] - Chlamydophila pneumoniae (st
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
C;Accession: D81593
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: D81593
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-952 <REA>
A;Cross-references: GB:AE002190; GB:AE002161; NID:g7189209; PIDN:AAF38142.1; PID:g718921
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: CP0285

  Query Match 20.9%; Score 1060.5; DB 2; Length 952;
  Best Local Similarity 30.6%; Pred. No. 4.4e-57;
  Matches 303; Conservative 164; Mismatches 447; Indels 77; Gaps 28;

```
Qy      6 FFFLIGNSLSGLAREVPSRIFLMPNSVPD------------PTKESLSNK------ISL 46
Db      8 FLFYLGNFTACMFGMTPAVYSLQTDSLEKFALERDEEFRTSFPLLDSLSTLTGFSPITTF 67

Qy     47 TGDTHNLT-NCYLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESG 105
Db     68 VGNRHNSSQDIVLSNYKSIDNILLLWTSAGGAVSCNNFL--LSNVEDHAFFSKNLAIGTG 125

Qy    106 GAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNPNAAVNKIREGGAIHAQ-NLYIN 164
Db    126 GAIACQG----ACTITKNRGPLIFFSNRGLN-------NASTGGETRGGAIACNGDFTIS 174

Qy    165 HNHDVVGFMKNFSYVRGGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYAGTSN 224
Db    175 QNQGTFYFVNNSVNNWGGALSTNGHCRIQSNRAPLLFFN------NTAPSGGGALRSENT 228
```

```
Qy    225 SFESNNCDLFFINNACCAGGAIFSPI-CSLTGNRGNIVFYNNRCFKNVETASSEASDGGA 283
Db    229 TISDNTRPIYFKNNCGNNGGAIQTSVTVAIKNNSGSVIFNNNTALSG--SINSGNGSGGA 286

Qy    284 IKVTTRLDVTGNRGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYID 343
Db    287 I-YTTNLSIDDNPGTILFNNNYCIRDGGAICTQFLTIKNSGHVYFTNN-QGNWGGALMLL 344

Qy    344 GTSNSKISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQN 403
Db    345 QDSTCLLFAEQGNIAFQNNEVFLTTFG----------RYNAIHCTPNS-NLQLGANKGYT 393

Qy    404 LIFYDPIEVSNAGVS-VSFNKEADQTGSVVFSGATVNSADFHQRNLQTKTPAPLTLSNGF 462
Db    394 TAFFDPIEHQHPTTNPLIFNPNANHQGTILFSSAYIPEASDYENNFISSSKNTSELRNGV 453

Qy    463 LCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNSAS-NASITLKHIGLNLSSILK 521
Db    454 LSIEDRAGWQFYKFTQKGGILKLGHAASIATTANSETPSTSVGSQVIINNLAINLPSILA 513

Qy    522 SGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTHALSSQPML 581
Db    514 KG-KAPTLWIRPLQSSAPFTEDNNPTITLSG-PLTLLNEENRDPYDSIDLSEPLQNIHLL 571

Qy    582 SISEASDNQLRSDDMDFSGLN-VPHYGWQGLWTWGWAKTQDPEPASSATITDPQKANRFH 640
Db    572 SLSDVTARHINTDNFHPESLNATEHYGYQGIWSPYWVET--ITTTNNASI---ETANTLY 626

Qy    641 RTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSD--HPFWGITGGG 698
Db    627 RALYANWTPLGYKVNPEYQGDLATTPLWQSFHTMFSLLRSYNRTGDSDIERPFLEIQGIA 686

Qy    699 LGMMVYQDPRENHPGFHMRSSGYS--AGMIAGQTHTFSLKFSQTYTKLNERYAKNNVSSK 756
Db    687 DGLFVHQNSIPGAPGFRIQSTGYSLQASSETSLHQKISLGFAQFFTRTKEIGSSNNVSAH 746

Qy    757 N--YSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVF 814
Db    747 NTVSSLYVELPW-FQEAF-ATSTVLAYGYGDHHLHSLHPSHQE-QAEGTCYSHTLAAAIG 803

Qy    815 FDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSF 874
Db    804 CSFPWQQKSYLHL--SPFVQAIAIRSHQTAFEEIGDNPRKFVSQKPFYNLTLPLGIQGKW 861

Qy    875 MNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQP 934
Db    862 QSKFHVPTEWTLELSYQPVLYQQNPQIGVTLLASGGSWDILGHNYVRNALGYKVHNQTAL 921

Qy    935 LSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
Db    922 FRSLDLFLDYQGSVSSSTSTHHLQAGSTLKF 952
```

RESULT 6
A72075
polymorphic outer membrane protein e family - Chlamydophila pneumoniae (strain CWL029
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
C;Accession: A72075
R;Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,
Nature Genet. 21, 385-389, 1999
A;Title: Comparative genomes of Clamydia pneumoniae and C. trachomatis.
A;Reference number: A72000; MUID:99206606; PMID:10192388
A;Accession: A72075
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-934 <ARN>
A;Cross-references: GB:AE001631; GB:AE001363; NID:g4376750; PIDN:AAD18609.1; PID:g437
A;Experimental source: strain CWL029
C;Genetics:
A;Gene: pmp_16

```
  Query Match             20.7%;  Score 1052;  DB 2;  Length 934;
  Best Local Similarity   31.2%;  Pred. No. 1.4e-56;
  Matches  296;  Conservative  158;  Mismatches  431;  Indels   64;  Gaps   27;

Qy      35 PTKESLSNK------ISLTGDTHNLT-NCYLDNLRYILAILQKTPNEGAAVTITDYLSFF 87
           |  :|||         :  |: || : :  | | : |  ||    : | ||:  ::|
Db      32 PLLDSLSTLTGFSPITTFVGNRHNSSQDIVLSNYKSIDNILLLWTSAGGAVSCNNFL--L 89

Qy      88 DTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNPNAAV 147
             ::  :|:|||   :||||          |    ||:|| :|           ||:
Db      90 SNVEDHAFFSKNLAIGTGGAIACQG----ACTITKNRGPLIFFSNRGLN-------NAST 138

Qy     148 NKIREGGAIHAQ-NLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSCFLFMDNIC 206
             ||||    :  |: |     |: |     |||:||     :  |::  || :
Db     139 GGETRGGAIACNGDFTISQNQGTFYFVNNSVNNWGGALSTNGHCRIQSNRAPLLFFN--- 195

Qy     207 IQTNTAGKGGAIYAGTSNSFESNNCDLFFINNACCAGGAIFSPI-CSLTGNRGNIVFYNN 265
           · |||  ||      : :   |   ::| ||    |||| : :  ::  | |:::| ||
Db     196 ---NTAPSGGGALRSENTTISDNTRPIYFKNNCGNNGGAIQTSVTVAIKNNSGSVIFNNN 252

Qy     266 RCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYAPVVTLVDNGP 325
             : :|    ||||  || | :  | | | |::|     ||||    :|: ::|
Db     253 TALSG--SINSGNGSGGAI-YTTNLSIDDNPGTILFNNNYCIRDGGAICTQFLTIKNSGH 309

Qy     326 TYFINNIANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAI 385
           || ||   | |||: :   |   : |::  | |  | |   |             | |||
Db     310 VYFTNN-QGNWGGALMLLQDSTCLLFAEQGNIAFQNNEVFLTTFG----------RYNAI 358

Qy     386 TVASSSGEILLGAGSSQNLIFYDPIEVSNAGVS-VSFNKEADQTGSVVFSGATVNSADFH 444
             :|  : |||       }:|||  :   : : || |:  |:::|| | :  |  :
Db     359 HCTPNS-NLQLGANKGYTTAFFDPIEHQHPTTNPLIFNPNANHQGTILFSSAYIPEASDY 417

Qy     445 QRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNSAS- 503
           : |  : :     | || | |||  |      :||| ||:: ||: | ::   |    | |
Db     418 ENNFISSSKNTSELRNGVLSIEDRAGWQFYKFTQKGGILKLGHAASIATTANSETPSTSV 477

Qy     504 NASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGN 563
           : : : :: :|| ||| | : | ||: |  :|  :| |   | :||   |:|:::
Db     478 GSQVIINNLAINLPSILAKG-KAPTLWIRPLQSSAPFTEDNNPTITLSG-PLTLLNEENR 535

Qy     564 SPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSGLN-VPHYGWQGLWTWGWAKTQDP 622
           ||:| ||:  | :  :||:|: :   : :|:     ||   |||:||:|:  | :|
Db     536 DPYDSIDLSEPLQNIHLLSLSDVTARHINTDNFHPESLNATEHYGYQGIWSPYWVET--I 593

Qy     623 EPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSA 682
           ::|:|   : ||  :| |   | | ||  :|::::  |    || :       |::
Db     594 TTTNNASI---ETANTLYRALYANWTPLGYKVNPEYQGDLATTPLWQSFHTMFSLLRSYN 650

Qy     683 ELTPSD--HPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYS--AGMIAGQTHTFSLKFS 738
             ||   ||  | |   |: |:|:     ||| ::|:|||  |           || |:
Db     651 RTGDSDIERPFLEIQGIADGLFVHQNSIPGAPGFRIQSTGYSLQASSETSLHQKISLGFA 710

Qy     739 QTYTKLNERYAKNNVSSKN--YSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQGE 796
           | :|:  |  : ||||: |   |   |: :  || |  |  |  | ||||: |  :   :
Db     711 QFFTRTKEIGSSNNVSAHNTVSSLYVELPW-FQEAF-ATSTVLAYGYGDHHLHSLHPSHQ 768

Qy     797 NLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSFS 856
             ::||  | |:  |:    | :     |: :||: |: | |  : |  |:|  || |
Db     769 E-QAEGTCYSHTLAAAIGCSFPWQQKSYLHL--SPFVQAIAIRSHQTAFEEIGDNPRKFV 825

Qy     857 TKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGSG 916
           :: |  |: :|:|::|  : :    |  ||:||:||||||:| | |   |||| | |   |
Db     826 SQKPFYNLTLPLGIQGKWQSKFHVPTEWTLELSYQPVLYQQNPQIGVTLLASGGSWDILG 885

Qy     917 SPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
             |:|: ||:  ||     | |   | |  ||||  ::|     |:|
Db     886 HNYVRNALGYKVHNQTALFRSLDLFLDYQGSVSSSTSTHHLQAGSTLKF 934


RESULT 7
G86548
polymorphic outer membrane protein E family [imported] - Chlamydophila pneumoniae (strai
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C;Accession: G86548
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.,
Nucleic Acids Res. 28, 2311-2314, 2000
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;Reference number: A86491; MUID:20330349; PMID:10871362
A;Accession: G86548
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-934 <STO>
A;Cross-references: GB:BA000008; NID:g8978837; PIDN:BAA98673.1; GSPDB:GN00142
A;Experimental source: strain J138
C;Genetics:
A;Gene: pmp_16

  Query Match             20.6%;  Score 1046;  DB 2;  Length 934;
  Best Local Similarity   31.1%;  Pred. No. 3.3e-56;
  Matches  295;  Conservative  158;  Mismatches  432;  Indels   64;  Gaps   27;

Qy      35 PTKESLSNK------ISLTGDTHNLT-NCYLDNLRYILAILQKTPNEGAAVTITDYLSFF 87
           |  :|||         :  |: || : :  | | : |  ||    : | ||:  ::|
Db      32 PLLDSLSTLTGFSPITTFVGNRHNSSQDIVLSNYKSIDNILLLWTSAGGAVSCNNFL--L 89

Qy      88 DTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNPNAAV 147
             ::  :|:|||   : |||          |    ||:|| :|           ||:
Db      90 SNVEDHAFFSKNLAIGTAGAIACQG----ACTITKNRGPLIFFSNRGLN-------NAST 138

Qy     148 NKIREGGAIHAQ-NLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSCFLFMDNIC 206
             ||||    :  |: |     |: |     |||:||     :  |::  || :
Db     139 GGETRGGAIACNGDFTISQNQGTFYFVNNSVNNWGGALSTNGHCRIQSNRAPLLFFN--- 195

Qy     207 IQTNTAGKGGAIYAGTSNSFESNNCDLFFINNACCAGGAIFSPI-CSLTGNRGNIVFYNN 265
           |||  ||      : :   |   ::| ||    |||| : :  ::  | |:::| ||
Db     196 ---NTAPSGGGALRSENTTISDNTRPIYFKNNCGNNGGAIQTSVTVAIKNNSGSVIFNNN 252

Qy     266 RCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYAPVVTLVDNGP 325
             : :|    ||||  || | :  | | | |::|     ||||    :|: ::|
Db     253 TALSG--SINSGNGSGGAI-YTTNLSIDDNPGTILFNNNYCIRDGGAICTQFLTIKNSGH 309

Qy     326 TYFINNIANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAI 385
           || ||   | |||: :   |   : |::  | |  | |   |             | |||
Db     310 VYFTNN-QGNWGGALMLLQDSTCLLFAEQGNIAFQNNEVFLTTFG----------RYNAI 358

Qy     386 TVASSSGEILLGAGSSQNLIFYDPIEVSNAGVS-VSFNKEADQTGSVVFSGATVNSADFH 444
             :|  : |||       |:||||  :   : : ||  |:  |:::|| | :  |  :
Db     359 HCTPNS-NLQLGANKGYTTAFFDPIEHQHPTTNPLIFNPNANHQGTILFSSAYIPEASDY 417

Qy     445 QRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNSAS- 503
           : |  : :     | || | |||  |      :||| ||:: ||: | ::   |    | |
Db     418 ENNFISSSKNTSELRNGVLSIEDRAGWQFYKFTQKGGILKLGHAASIATTANSETPSTSV 477

Qy     504 NASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGN 563
           : : : :: :|| ||| | : | ||: |  :|  :| |   | :||   |:|:::
Db     478 GSQVIINNLAINLPSILAKG-KAPTLWIRPLQSSAPFTEDNNPTITLSG-PLTLLNEENR 535

Qy     564 SPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSGLN-VPHYGWQGLWTWGWAKTQDP 622
           ||:| ||:  | :  :||:|: :   : :|:     ||   |||:||:|:  | :|
Db     536 DPYDSIDLSEPLQNIHLLSLSDVTARHINTDNFHPESLNATEHYGYQGIWSPYWVET--I 593

Qy     623 EPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSA 682
           ::|:|   : ||  :| |   | | ||  :|::::  |    || :       |::
Db     594 TTTNNASI---ETANTLYRALYANWTPLGYKVNPEYQGDLATTPLWQSFHTMFSLLRSYN 650

Qy     683 ELTPSD--HPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYS--AGMIAGQTHTFSLKFS 738
             ||   ||  | |   |: |:|:     ||| ::|:|||  |           || |:
Db     651 RTGDSDIERPFLEIQGIADGLFVHQNSIPGAPGFRIQSTGYSLQASSETSLHQKISLGFA 710

Qy     739 QTYTKLNERYAKNNVSSKN--YSCQGEMLFSLQEGFLLTKLVGLYSYGDHNCHHFYTQGE 796
           | :|:  |  : ||||: |   |   |: :  || |  |  |  | ||||: |  :   :
Db     711 QFFTRTKEIGSSNNVSAHNTVSSLYVELPW-FQEAF-ATSTVLAYGYGDHHLHSLHPSHQ 768

Qy     797 NLTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSFS 856
             ::||  | |:  |:    | :     |: :||: |: | |  : | |:|  || |
```

```
Db    769 E-QAEGTCYSHTLAAAIGCSFPWQQKSYLHL--SPFVQAIAIRSHQTAFEEIGDNPRKFV 825

Qy    857 TKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGSG 916
          :: |  |: :|:|::| : :    |  ||:||:||||||:| | |   |||| | |   |
Db    826 SQKPFYNLTLPLGIQGKWQSKFHVPTEWTLELSYQPVLYQQNPQIGVTLLASGGSWDILG 885

Qy    917 SPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
            |:|: ||:  ||     | |   | |  ||||  ::|     |:|
Db    886 HNYVRNALGYKVHNQTALFRSLDLFLDYQGSVSSSTSTHHLQAGSTLKF 934


RESULT 8
C86549
polymorphic outer membrane protein E/F family [imported] - Chlamydophila pneumoniae (str
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C;Accession: C86549
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
Nucleic Acids Res. 28, 2311-2314, 2000
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;Reference number: A86491; MUID:20330349; PMID:10871362
A;Accession: C86549
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-946 <STO>
A;Cross-references: GB:BA000008; NID:g8978841; PIDN:BAA98677.1; GSPDB:GN00142
A;Experimental source: strain J138
C;Genetics:
A;Gene: pmp_18

  Query Match             18.7%;  Score 951.5;  DB 2;  Length 946;
  Best Local Similarity   29.1%;  Pred. No. 2.1e-50;
  Matches  281;  Conservative  168;  Mismatches  416;  Indels  101;  Gaps   34;

Qy     35 PTKESLSNKISLTGDTHNLT----------NCYLDNLRYILAILQKTPNEGAAVTITDYL 84
          |  ::|:|   :|  :|  |          :  ||:   | :  :    :| |::    |
Db     47 PLIDTLTN---MTPYSHRATLFGVRDDTNQDIVLDHQNSIESWFENFSQDGGALSCKS-L 102

Qy     85 SFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNPN 144
          :  :| |  | |  :   :  ||: | : |    :: :  | :||  |          | ||
Db    103 AITNT-KNQILFLNSFAIKRAGAM-YVNGN---FDLSENHGSIIFSGNL-------SFPN 150

Qy    145 AA--VNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSCFLFM 202
          |:   :    |  : ::|: |: |     |: | :   ||||  |   :       |  |
Db    151 ASNFADTCTGGAVLCSKNVTISKNQGTAYFINNKAKSSGGAIQAAIINIKDNTGPCLFF- 209

Qy    203 DNICIQTNTAG--KGGAIYAGTSNSFESNNCDLFFINNACCAGGAI-FSPICSLTGNRGN 259
                | ||   |||::|      |:|:  ::|:||    ||||     | || | |:
Db    210 ------NNAAGGTAGGALFANACR-IENNSQPIYFLNNQSGLGGAIRVHQECILTKNTGS 262

Qy    260 IVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYAPVVT 319
          ::| ||  |      |:  | ||||   :   :  | |   | :|     ||||   :|
Db    263 VIFNNN--FAMEADISANHSSGGAIYCIS-CSIKDNPGIAAFDNNTAARDGGAICTQSLT 319

Qy    320 LVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTSTSANP 379
          : |:|| || ||    |||| :       : ||:  |||  |     | :|    | |
Db    320 IQDSGPVYFTNN-QGTWGGAIMLRQDGACTLFADQGDIIFYNNRHFKDTFSN--HVSVNC 376

Qy    380 PRRNAITVASSSGEILLGAGSSQNLIFYDPI----EVSNAGVSVSFNKEADQTGSVVFSG 435
          |  ::|| :| |          :  |||||     : |:     ||   :  |:::||
Db    377 TRNVSLTVGASQG---------HSATFYDPILQRYTIQNS--IQKFNPNPEHLGTILFSS 425

Qy    436 A-----TVNSADF--HQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNG 488
          |     :  ||  | ||        : | || | :|| |:  | :| | || : ||:
Db    426 AYIPDTSTSRDDFISHFRN-------HIGLYNGTLALEDRAEWKVYKFDQFGGTLRLGSR 478

Qy    489 AVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATF 548
          || |       :|:  : | : :: :|| |||      | ||: || :|  |: |
Db    479 AVFSTTDEEQSSSSVGSVININNLAINLPSIL-GNRVAPKLWIRPTGSSAPYSEDNNPII 537

Qy    549 SLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSGLN-VPHYG 607
          :||   |||:||    ||:: ||   ::  |:| : :  :  : :|:    |||   |||
Db    538 NLSG-PLSLLDDENLDPYDTADLAQPIAEVPLLYLLDVTAKHINTDNFYPEGLNTTQHYG 596

Qy    608 WQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTL 667
          :||:|:  | :|      ||         |  || |   | | ||  :|:::  :  :
Db    597 YQGVWSPYWIETITTSDTSS-----EDTVNTLHRQLYGDWTPTGYKVNPENKGDIALSAF 651

Qy    668 W---GNMLLATESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAG 724
          |    |:             :: |:       :|    : |:|:   :  |||| ::||| |
Db    652 WQSFHNLFATLRYQTQQGQIAPT------ASGEATRLFVHQNSNNDAKGFHMEATGYSLG 705

Qy    725 MIA--GQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFS-LQEGFLLTKLVGLY 781
            :     |:| : ||| :: | | :: |:|:|   :   ::    ||| |  :  :  |
Db    706 TTSNTASNHSFGVNFSQLFSNLYESHSDNSVASHTTTVALQINNPWLQERFSTSASLA-Y 764

Qy    782 SYGDHNCHHFYTQGEN--LTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIY 839
          || :  ||    | :  : ::|   | |:| |:   | :: : |  :   ||: |: :
Db    765 SYSN---HHIKASGYSGKIQTEGKCYSTTLGAALSCSLSLQ-WRSRPLHFTPFIQAIAVR 820

Qy    840 SSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEP 899
          |: : | | |   | ||   || |: ||:|:: :: :  : |  | :|||||||||:| |
Db    821 SNQTAFQESGDKARKFSVHKPLYNLTVPLGIQSAWESKFRLPTYWNIELAYQPVLYQQNP 880

Qy    900 GIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNG 959
            :  | :|   |  ||:  :|:|:::|   |      |::   | |  ||||  :||:
Db    881 EVNVSLESSGSSWLLSGTTLARNAIAFKGRNQIFIFPKLSVFLDYQGSVSSSTTTHYLHA 940

Qy    960 EIALRF 965
            :|
Db    941 GTTFKF 946


RESULT 9
D81594
polymorphic membrane protein E/F family CP0283 [imported] - Chlamydophila pneumoniae
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
C;Accession: D81594
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: D81594
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-946 <REA>
A;Cross-references: GB:AE002189; GB:AE002161; NID:g7189205; PIDN:AAF38141.1; PID:g718
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: CP0283

  Query Match             18.7%;  Score 951.5;  DB 2;  Length 946;
  Best Local Similarity   29.1%;  Pred. No. 2.1e-50;
  Matches  281;  Conservative  168;  Mismatches  416;  Indels  101;  Gaps   34;

Qy     35 PTKESLSNKISLTGDTHNLT----------NCYLDNLRYILAILQKTPNEGAAVTITDYL 84
          |  ::|:|   :|  :|  |          :  ||:   | :  :    :| |::    |
Db     47 PLIDTLTN---MTPYSHRATLFGVRDDTNQDIVLDHQNSIESWFENFSQDGGALSCKS-L 102

Qy     85 SFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNPN 144
          :  :| |  | |  :   :  ||: | : |    :: :  | :||  |          | ||
Db    103 AITNT-KNQILFLNSFAIKRAGAM-YVNGN---FDLSENHGSIIFSGNL-------SFPN 150

Qy    145 AA--VNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSCFLFM 202
          |:   :    |  : ::|: |: |     |: | :   ||||  |   :       |  |
Db    151 ASNFADTCTGGAVLCSKNVTISKNQGTAYFINNKAKSSGGAIQAAIINIKDNTGPCLFF- 209

Qy    203 DNICIQTNTAG--KGGAIYAGTSNSFESNNCDLFFINNACCAGGAI-FSPICSLTGNRGN 259
                | ||   |||::|      |:|:  ::|:||    ||||     | || | |:
Db    210 ------NNAAGGTAGGALFANACR-IENNSQPIYFLNNQSGLGGAIRVHQECILTKNTGS 262

Qy    260 IVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYAPVVT 319
          ::| ||  |      |:  | ||||   :   :  | |   | :|     ||||   :|
Db    263 VIFNNN--FAMEADISANHSSGGAIYCIS-CSIKDNPGIAAFDNNTAARDGGAICTQSLT 319
```

```
Qy     320 LVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTSTSANP 379
Db     320 IQDSGPVYFTNN-QGTWGGAIMLRQDGACTLFADQGDIIFYNNRHFKDTFSN--HVSVNC 376

Qy     380 PRRNAITVASSSGEILLGAGSSQNLIFYDPI----EVSNAGVSVSFNKEADQTGSVVFSG 435
Db     377 TRNVSLTVGASQG---------HSATFYDPILQRYTIQNS--IQKFNPNPEHLGTILFSS 425

Qy     436 A-----TVNSADF--HQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNG 488
Db     426 AYIPDTSTSRDDFISHFRN-------HIGLYNGTLALEDRAEWKVYKFDQFGGTLRLGSR 478

Qy     489 AVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATF 548
Db     479 AVFSTTDEEQSSSSVGSVININNLAINLPSIL-GNRVAPKLWIRPTGSSAPYSEDNNPII 537

Qy     549 SLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSGLN-VPHYG 607
Db     538 NLSG-PLSLLDDENLDPYDTADLAQPIAEVPLLYLLDVTAKHINTDNFYPEGLNTTQHYG 596

Qy     608 WQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTL 667
Db     597 YQGVWSPYWIETITTSDTSS-----EDTVNTLHRQLYGDWTPTGYKVNPENKGDIALSAF 651

Qy     668 W---GNMLLATESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAG 724
Db     652 WQSFHNLFATLRYQTQQGQIAPT------ASGEATRLFVHQNSNNDAKGFHMEATGYSLG 705

Qy     725 MIA--GQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFS-LQEGFLLTKLVGLY 781
Db     706 TTSNTASNHSFGVNFSQLFSNLYESHSDNSVASHTTTVALQINNPWLQERFSTSASLA-Y 764

Qy     782 SYGDHNCHHFYTQGEN--LTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIY 839
Db     765 SYSN---HHIKASGYSGKIQTEGKCYSTTLGAALSCSLSLQ-WRSRPLHFTPFIQAIAVR 820

Qy     840 SSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEP 899
Db     821 SNQTAFQESGDKARKFSVHKPLYNLTVPLGIQSAWESKFRLPTYWNIELAYQPVLYQQNP 880

Qy     900 GIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNG 959
Db     881 EVNVSLESSGSSWLLSGTTLARNAIAFKGRNQIFIFPKLSVFLDYQGSVSSSTTTHYLHA 940

Qy     960 EIALRF 965
Db     941 GTTFKF 946
```

RESULT 10
C72075
polymorphic outer membrane protein e/f family - Chlamydophila pneumoniae (strain CWL029)
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
C;Accession: C72075
R;Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A;Title: Comparative genomes of Clamydia pneumoniae and C. trachomatis.
A;Reference number: A72000; MUID:99206606; PMID:10192388
A;Accession: C72075
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-946 <ARN>
A;Cross-references: GB:AE001631; GB:AE001363; NID:g4376750; PIDN:AAD18610.1; PID:g437675
A;Experimental source: strain CWL029
C;Genetics:
A;Gene: pmp_18

Query Match 18.6%; Score 948.5; DB 2; Length 946;
Best Local Similarity 29.1%; Pred. No. 3.2e-50;
Matches 281; Conservative 167; Mismatches 417; Indels 101; Gaps 34;

```
Qy      35 PTKESLSNKISLTGDTHNLT----------NCYLDNLRYILAILQKTPNEGAAVTITDYL 84
Db      47 PLIDTLTN---MTPYSHRATLFGVRDDTNQDIVLDHQNSIESWFENFSQDGGALSCKS-L 102

Qy      85 SFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNPN 144
Db     103 AITNT-KNQILFLNSFAIKRAGAM-YVNGN---FDLSENHGSIIFSGNL-------SFPN 150

Qy     145 AA--VNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSCFLFM 202
Db     151 ASNFADTCTGGAVLCSKNVTISKNQGTAYFINNKAKSSGGAIQAAIINIKDNTGPCLFF- 209

Qy     203 DNICIQTNTAG--KGGAIYAGTSNSFESNNCDLFFINNACCAGGAI-FSPICSLTGNRGN 259
Db     210 ------NNAAGGTAGGALFANACR-IENNSQPIYFLNNQSGLGGAIRVHQECILTKNTGS 262

Qy     260 IVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYAPVVT 319
Db     263 VIFNNN--FAMEADISANHSSGGAIYCIS-CSIKDNPGIAAFDNNTAARDGGAICTQSLT 319

Qy     320 LVDNGPTYFINNIANNKGGAIYIDGTSNSKISADRHAIIFNENIVTNVTNANGTSTSANP 379
Db     320 IQDSGPVYFTNN-QGTWGGAIMLRQDGACTLFADQGDIIFYNNRHFKDTFSN--HVSVNC 376

Qy     380 PRRNAITVASSSGEILLGAGSSQNLIFYDPI----EVSNAGVSVSFNKEADQTGSVVFSG 435
Db     377 TRNVSLTVGASQG---------HSATFYDPILQRYTIQNS--IQKFNPNPEHLGTILFSS 425

Qy     436 -----ATVNSADF--HQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNG 488
Db     426 TYIPDTSTSRDDFISHFRN-------HIGLYNGTLALEDRAEWKVYKFDQFGGTLRLGSR 478

Qy     489 AVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATF 548
Db     479 AVFSTTDEEQSSSSVGSVININNLAINLPSIL-GNRVAPKLWIRPTGSSAPYSEDNNPII 537

Qy     549 SLSDVKLSLIDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSDDMDFSGLN-VPHYG 607
Db     538 NLSG-PLSLLDDENLDPYDTADLAQPIAEVPLLYLLDVTAKHINTDNFYPEGLNTTQHYG 596

Qy     608 WQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTL 667
Db     597 YQGVWSPYWIETITTSDTSS-----EDTVNTLHRQLYGDWTPTGYKVNPENKGDIALSAF 651

Qy     668 W---GNMLLATESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFHMRSSGYSAG 724
Db     652 WQSFHNLFATLRYQTQQGQIAPT------ASGEATRLFVHQNSNNDAKGFHMEATGYSLG 705

Qy     725 MIA--GQTHTFSLKFSQTYTKLNERYAKNNVSSKNYSCQGEMLFS-LQEGFLLTKLVGLY 781
Db     706 TTSNTASNHSFGVNFSQLFSNLYESHSDNSVASHTTTVALQINNPWLQERFSTSASLA-Y 764

Qy     782 SYGDHNCHHFYTQGEN--LTSQGTFRSQTMGGAVFFDLPMKPFGSTHILTAPFLGALGIY 839
Db     765 SYSN---HHIKASGYSGKIQTEGKCYSTTLGAALSCSLSLQ-WRSRPLHFTPFIQAIAVR 820

Qy     840 SSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVLYRQEP 899
Db     821 SNQTAFQESGDKARKFSVHKPLYNLTVPLGIQSAWESKFRLPTYWNIELAYQPVLYQQNP 880

Qy     900 GIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNG 959
Db     881 EINVSLESSGSSWLLSGTTLARNAIAFKGRNQIFIFPKLSVFLDYQGSVSSSTTTHYLHA 940

Qy     960 EIALRF 965
Db     941 GTTFKF 946
```

RESULT 11
F71460
probable outer membrane protein F - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C;Species: Chlamydia trachomatis

```
C;Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
C;Accession: F71460
R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,
Science 282, 754-759, 1998
A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trac
A;Reference number: A71570; MUID:99000809; PMID:9784136
A;Accession: F71460
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1034 <ARN>
A;Cross-references: GB:AE001360; GB:AE001273; NID:g3329342; PIDN:AAC68468.1; PID:g332934
A;Experimental source: serotype D, strain UW-3/Cx
C;Genetics:
A;Gene: pmpF

  Query Match            17.0%;  Score 866.5;  DB 2;  Length 1034;
  Best Local Similarity  28.1%;  Pred. No. 3.9e-45;
  Matches  293;  Conservative  149;  Mismatches  406;  Indels  193;  Gaps   36;

Qy      55 NCYLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAI------ 108
Db      57 NIVLSNLQ---------SNGTGACTISG-----NTQTQIFSNSVNTTADSGGAFDMVTTS 102

Qy     109 ------------GYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNPNAAV------NK 149
Db     103 FTASDNANLLFCNNYCTHNKGGGAIRSG-GPIRFLNNQDVLFYNNISAGAKYVGTGDHNE 161

Qy     150 IREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSCFLFMDNICI-- 207
Db     162 KNRGGALYATTITLTGNR-TLAFINNMSGDCGGAISADTQISITDTVKGILFENNHTLNH 220

Qy     208 ----QTNTAGKGGAIYAGTS-NSFESNNCDLFFINNACCAGGAIFSPICSLTGNRGNIVF 262
Db     221 IPYTQAENMARGGAICSRRDLCSISNNSGPIVFNYNQGGKGGAISATRCVIDNNKERIIF 280

Qy     263 YNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIFFSDNITKNYGGAIYAPVVTLVD 322
Db     281 SNNSSLG--WSQSSSASNGGAIQTTQGFTLRNNKGSIYFDSNTATHAGGAINCGYIDIRD 338

Qy     323 NGPTYFINNIANNKGGAIYIDGTSNSKISADRHA----------IIFNENIVTNVTNANG 372
Db     339 NGPVYFLNNSA--AWGAAF-------NLSKPRSATNYIHTGTGDIVFNNNVVFTLDGNLL 389

Qy     373 TSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDPIE-------VSNAGVS------V 419
Db     390 GKRKLFHINNNEITPYTLS----LGAKKDTRIYFYDLFQWERVKENTSNNPPSPTSRNTI 445

Qy     420 SFNKEADQTGSVVFSGATVNSADFHQRNLQTK-----TPAPLTLSNGFLCIEDHAQLTVN 474
Db     446 TVNPETEFSGAVVFS---YNQMSSDIRTLMGKEHNYIKEAPTTLKFGTLAIEDDAELEIF 502

Qy     475 R--FTQT-GGVVSLGNGAVLSCYKNGAGNSASNASITLKHIGLNLSSILKSGAEIPLLWV 531
Db     503 NIPFTQNPTSLLALGSGATLTVGKHG--------KLNITNLGVILPIILKEGKSPPCIRV 554

Qy     532 EPTN-NSNNYTADT-AATFSLS------DVKLSLIDDYGNSPYESTDLTHALSSQPMLSI 583
Db     555 NPQDMTQNTGTGQTPSSTSSISTPMIIFNGRLSIVDENYESVYDSMDLSRGKAEQLILSI 614

Qy     584 SEASDNQLRSD---DMDFSGLNVPHYGWQGLWTWGWAKT--------QDPEPASSATITD 632
Db     615 ETTNDGQLDSNWQSSLNTSLLSPPHYGYQGLWTPNWITTTYTITLNNNSSAPTSATSIAE 674

Qy     633 PQKANR----------------------------------FHRTLLLTWLPAGYVPSPKH 658
Db     675 QKKTSETFTPSNTTTASIPNIKASAGSGSGSASNSGEVTITKHTLVVNWAPVGYIVDPIR 734

Qy     659 RSPLIANTL---WGNMLLATESLKNSAELTPSDHPFWGITGGGLGMMVYQDPRENHPGFH 715
Db     735 RGDLIANSLVHSGRNMTMGLRSLL-------PDNSWFALQGAATTLFTKQQKRLSYHGYS 787

Qy     716 MRSSGYSAGMIA--GQTHTFSLKFSQTYTKLNERYAKNNVSSKNY---SCQGEMLFSLQE 770

Db     788 SASKGYTVSSQASGAHGHKFLLSFSQSSDKMKEKETNNRLSSRYYLSALCFEHPMFD--- 844

Qy     771 GFLLTKLVG--LYSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDL----PMKPFGS 824
Db     845 ---RIALIGAAACNYGTHNMRSFY--GTKKSSKGKFHSTTLGASLRCELRDSMPLR---- 895

Qy     825 THILTAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAW 884
Db     896 -SIMLTPFAQALFSRTEPASIRESGDLARLFTLEQAHTAVVSPIGIKGAYSSDTWPTLSW 954

Qy     885 TVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQY 944
Db     955 EMELAYQPTLYWKRPLLNTLLIQNNGSWVTTNTPLAKHSF-YGRGSHSLKFSHLKLFANY 1013

Qy     945 HGFYSSSTFCNYLNGEIALRF 965
Db    1014 QAEVATSTVSHYINAGGALVF 1034


RESULT 12
G81722
polymorphic membrane protein E/F family TC0262 [imported] - Chlamydia muridarum (stra
C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
C;Accession: G81722
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: G81722
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1025 <TET>
A;Cross-references: GB:AE002293; GB:AE002160; NID:g7190298; PIDN:AAF39131.1; PID:g719
A;Experimental source: strain Nigg (MoPn)
C;Genetics:
A;Gene: TC0262

  Query Match            15.6%;  Score 793;  DB 2;  Length 1025;
  Best Local Similarity  27.0%;  Pred. No. 1.3e-40;
  Matches  269;  Conservative  158;  Mismatches  419;  Indels  152;  Gaps   35;

Qy      68 LQKTPNEGAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGYASPNSPTVEIRDTIGPV 127
Db      80 LSTTNSSGAFV--ESMTSFTAIDNADLFFCNNYCTHQGG--GGAINATGLISFKNNQNIL 135

Qy     128 IFENNTCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTA 187
Db     136 FYNNTTIGTQFTGVALRTERNR---GGALYGSSIELINNHS-LNFINNTSGDMGGAVSTI 191

Qy     188 NTFVVSENQSCFLFMDNICIQ--TNT----AGKGGAIYAGTSNSFESNNCDLFFINNACC 241
Db     192 QNLVIKNTSGIVAFENNHTTDHIPNTFATILARGGAVGCQGACEISHNTGPVVFNSNYGG 251

Qy     242 AGGAIFS-PICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTTRLDVTGNRGRIF 300
Db     252 YGGAISTGGQCIFRDNKDKLIFINNSALGWHNT--SAQGNGAVISAGGEFGLLNNKGPIY 309

Qy     301 FSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNKGGAIYIDGT-SNSKISADRHAIIF 359
Db     310 FENNNASYIAGAISCNNLNFQENGPIYFLNNSA-LYGGAFHLFASPAANYIHTGSGDIIF 368

Qy     360 NENIVTNVTNAN----------GTSTSANPPRRNAITVASSSGEILLGAGSSQNLIFYDP 409
Db     369 NNNTELSTTGMSAGLRKLFYIPGTTNN------NPITLS-------LGAKKDTRIYFYDL 415

Qy     410 IE----------VSNAGVSVSFNKEADQTGSVVFSGATVNSADFHQRNLQTKT-----PA 454
Db     416 FQWGGLKKANTPPENSPHTVTINPSDEFSGAVVFSYKNI-SSDLQAHMIASKTHNQIKDS 474

Qy     455 PLTLSNGFLCIEDHAQ-------LTVNRFTQTGGVVSLGNGAVLSCYKNGAGNSASNASI 507
```

```
Db    475 PTTLKFGTMSIENGAEFEFFNGPLT----QESTSLLALGQDSILTVGK--------DASL 522

Qy    508 TLKHIGLNLSSILK-SGAEIPLLWVEP---TNNSNNYTADTAATFSLSDVKL------SL 557
Db    523 TITHLGIILPGLLNDQGTTAPRIRVNPQDMTQNTNSNQA-PVSTENVATQKIFFSGLVSL 581

Qy    558 IDDYGNSPYESTDLTHALSSQPMLSISEASDNQLRSD---DMDFSGLNVPHYGWQGLWTW 614
Db    582 VDENYESVYDSCDLSRGKANQPILHIETTNDAQLSNDWKNTLNTSLYSLPHYGYQGLWTS 641

Qy    615 GWAKTQDPEPASSATITDPQKANR---------------------------------- 638
Db    642 NWMTT--TRTVSLTNSTETQTANNSIQEQKNTSETFDSNSTTTAKIPSIRASTGGTTPLA 699

Qy    639 ------FHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSDHPFW 692
Db    700 TTDVTVTRHSLVVSWTPIGYIADPARRGDLIANNLVSSGRNTTLYLRS---LLPDD-SWF 755

Qy    693 GITGGGLGMMVYQDPRENHPGFHMRSSGY--SAGMIAGQTHTFSLKFSQTYTKLNERYAK 750
Db    756 ALQGSAATLFTKQQKRLDYHGYSSASKGYAISSQASGAHGHKFLFSFSQSSDTMKEKRTN 815

Qy    751 NNVSSKNYSCQGEMLFSLQEGFLLTKLVG--LYSYGDHNCHHFYTQGENLTSQGTFRSQT 808
Db    816 NKISSRYYL---SALCFEQPMFDRIALIGAAAYNYGTHKTYNFY--GTKKFSKGNFHSTT 870

Qy    809 MGGAVFFDL-PMKPFGSTHILTAPFLGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVP 867
Db    871 LGGSLRCELRDSMPFQS--IMLTPFIQALISRTEPASIQEQGDLARLFSLKQPHTAVVSP 928

Qy    868 IGVKGSFMNATQRPQAWTVELAYQPVLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYK 927
Db    929 IGIKGVYSSNKWPTVSCEMEVAYQPTLYWKRPILNTVLIKNNGSWETTNTPLAKHSF-YG 987

Qy    928 ISQQTQPLSWLTLHFQYHGFYSSSTFCNYLNGEIALRF 965
Db    988 RGSSSLKFSYLKLFANYQAQVATSTVSHYMNAGGALVF 1025
```

RESULT 13
G71460
probable outer membrane protein G - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C;Species: Chlamydia trachomatis
C;Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
C;Accession: G71460
R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,
Science 282, 754-759, 1998
A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trac
A;Reference number: A71570; MUID:99000809; PMID:9784136
A;Accession: G71460
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1013 <ARN>
A;Cross-references: GB:AE001360; GB:AE001273; NID:g3329342; PIDN:AAC68469.1; PID:g332934
A;Experimental source: serotype D, strain UW-3/Cx
C;Genetics:
A;Gene: pmpG

  Query Match 12.9%; Score 655.5; DB 2; Length 1013;
  Best Local Similarity 25.7%; Pred. No. 3.4e-32;
  Matches 261; Conservative 152; Mismatches 388; Indels 213; Gaps 43;

```
Qy     39 SLSNKISLTGDTHNLTNCYLDNLRYILAILQKTPNEGAAVTITDYLSFFDTQKEGIYFAK 98
Db     84 NLLGSFTVLGRGHSLT---FENIR--------TSTNGAALS--------DSANSGLF--- 121

Qy     99 NLTPESGGAIGYASPNSPTVEIRDTIGPVIFENNTCCRPFTSSNP-NAAV---------- 147
Db    122 --TIEGFKELSFSNCNSLLAVL-----PAATTNNGSQTPTTTSTPSNGTIYSKTDLLLLN 174

Qy    148 --------NKIR-EGGAIHAQNLYINHNHDVVGFMKNFSYVRGGAISTANTFVVSENQSC 198
Db    175 NEKFSFYSNLVSGDGGAIDAKSLTVQGISKLCVFQENTAQADGGACQVVTSFSAMANEAP 234

Qy    199 FLFMDNICIQTNTAG-KGGAIYA------GTSNSFE--------SNNCDLFFINNACCAG 243
Db    235 IAFIANV------AGVRGGGIAAVQDGQQGVSSSTSTEDPVVSFSRNTAVEFDGNVARVG 288

Qy    244 GAIFSPICSLTGNRGNIVFYNN--RCFKN-----------------VETASSEASDGGAI 284
Db    289 GGIYS--------YGNVAFLNNGKTLFLNNVASPVYIAAEQPTNGQASNTSDNYGDGGAI 340

Qy    285 KVTTRLDVTGN----------RGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIAN 334
Db    341 FCKNGAQAAGSNNSGSVSFDGEGVVFFSSNVAAGKGGAIYAKKLSVANCGPVQFLGNIAN 400

Qy    335 NKGGAIYIDGTSNSKISADRHAIIFNENI----VTNVTNANGTSTSANPPRRNAITVASS 390
Db    401 D-GGAIYLGESGELSLSADYGDIIFDGNLKRTAKENAADVNGVTVSS-----QAISMGSG 454

Qy    391 SGEILLGAGSSQNLIFYDPIEVSN-------AGVSVSFNKEADQTGSVVFSGATVNSADF 443
Db    455 GKITTLRAKAGHQILFNDPIEMANGNNQPAQSSEPLKINDGEGYTGDIVF--ANGNSTLY 512

Qy    444 HQRNLQTKTPAPLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNS-- 501
Db    513 QN----------VTIEQGRIVLREKAKLSVNSLSQTGGSLYMEAGSTLDFVTPQPPQQPP 562

Qy    502 ASNASITLKHIGLNLSSILKSGAEIPLLWVEPTN--NSNNYTADTAATFSLSDVKLS--- 556
Db    563 AANQLITLSNLHLSLSSLLANNA----VTNPPTNPPAQDSHPAIIGST-TAGSVTISGPI 617

Qy    557 LIDDYGNSPYESTD----------LTHALSSQPMLSISEASDNQLRSDDMDFSGLNVPHY 606
Db    618 FFEDLDDTAYDRYDWLGSNQKIDVLKLQLGTQP--SANAPSDLTL--------GNEMPKY 667

Qy    607 GWQGLWTWGWAKTQDPEPASSATITDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANT 666
Db    668 GYQGSWKLAW----------DPNTANNGPYTLKATWTKTGYNPGPERVASLVPNS 712

Qy    667 LWGNMLLATESLKNSAELTPSDHPFW-GITGGGLGMMVYQDPRENHPGFHMRSSGYSAGM 725
Db    713 LWGS-ILDIRSAHSAIQASVDGRSYCRGLWVSGVSNFFYHDRDALGQGYRYISGGYSLG- 770

Qy    726 IAGQTHTFSLKFSQTYTKLNERYAKNNV--SSKNYSCQGEMLFSLQEGFLLTKLVGLYSY 783
Db    771 --ANSYFGSSMFGLAFTEVFGR-SKDYVVCRSNHHACIGSVYLSTKQA-----LCGSYLF 822

Qy    784 GDHNCHHFYTQG-ENLTSQGTFRSQT--------MGGAVFFDLPMKPFGSTHILT--APF 832
Db    823 GDAFIRASYGFGNQHMKTSYTFAEESDVRWDNNCLVGEIGVGLPIVITPSKLYLNELRPF 882

Qy    833 LGALGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQP 892
Db    883 VQAEFSYADHESFTEEGDQARAFRS-GHLMNLSVPVGVKFDRCSST-HPNKYSFMGAYIC 940

Qy    893 VLYRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHG 946
Db    941 DAYRTISGTQTTLLSHQETWTTDAFHLARHGVIVRGSMYAS----LTSNIEVYG 990
```

RESULT 14
C72078
polymorphic outer membrane protein g family - Chlamydophila pneumoniae (strain CWL029
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 15-Jun-2001
C;Accession: C72078
R;Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,
Nature Genet. 21, 385-389, 1999
A;Title: Comparative genomes of Clamydia pneumoniae and C. trachomatis.
A;Reference number: A72000; MUID:99206606; PMID:10192388
A;Accession: C72078
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-936 <ARN>
A;Cross-references: GB:AE001627; GB:AE001363; NID:g4376721; PIDN:AAD18589.1; PID:g437

A;Experimental source: strain CWL029
C;Genetics:
A;Gene: pmp_7
C;Superfamily: Chlamydophila pneumoniae polymorphic outer membrane protein G

Query Match 12.7%; Score 645; DB 2; Length 936;
Best Local Similarity 27.0%; Pred. No. 1.3e-31;
Matches 262; Conservative 146; Mismatches 392; Indels 172; Gaps 44;

```
Qy      75 GAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGY------------ASPNSPTVEIR 121
Db      54 GTTYSLLSDVSFQNAGALGIPLASGCFLEAGGDLTFQGNQHALKFAFINAGSSAGTVAST 113

Qy     122 DTIGPVIFENN-------TCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMK 174
Db     114 SAADKNLLFNDFSRLSIISCPSLLLSPTGQCALKSV--------GNLSLTGNSQII-FTQ 164

Qy     175 NFSYVRGGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYA-GTSNSFESNNCDL 233
Db     165 NFSSDNGGVINTKN-FLLSGTSQFASFSRN---QAFTGKQGGVVYATGTITIENSPGIVS 220

Qy     234 FFINNACCAGGAIFS-PICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTT---R 289
Db     221 FSQNLAKGSGGALYSTDNCSITDN-FQVIFDGNSAWEAAQ------AQGGAICCTTTDKT 273

Qy     290 LDVTGNRGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNK-----GGAIYIDG 344
Db     274 VTLTGNK-NLSFTNNTALTYGGAISGLKVSISAGGPTLFQSNISGSSAGQGGGGAINIAS 332

Qy     345 TSNSKISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNL 404
Db     333 AGELALSATSGDITFNNNQVTN----GSTST------RNAINIIDTAKVTSIRAATGQSI 382

Qy     405 IFYDPIEVSNAGVSVS----------FNKEADQTGSVVFSGATVNSAD-FHQRNLQTKTP 453
Db     383 YFYDPI--TNPGTAASTDTLNLNLADANSEIEYGGAIVFSGEKLSPTEKAIAANVTSTIR 440

Qy     454 APLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNSASNASITLKHIG 513
Db     441 QPAVLARGDLVLRDGVTVTFKDLTQSPGSRILMDG--------GTTLSAKEANLSLNGLA 492

Qy     514 LNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTH 573
Db     493 VNLSSL------------DGTNKAALKTEAADKNISLSGT-IALIDTEG-SFYENHNLKS 538

Qy     574 ALSSQPMLSISEASDN-QLRSDDMDFSGLNVP--HYGWQGLWTWGWAKTQDPEPASSATI 630
Db     539 A-STYPLLELTTAGANGTITLGALSTLTLQEPETHYGYQGNWQLSWAN------ATSSKI 591

Qy     631 TDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSDHP 690
Db     592 GS------------INWTRTGYIPSPERKSNLPLNSLWGN-FIDIRSINQLIETKSSGEP 638

Qy     691 F----WGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFSLKFSQTYTKLNE 746
Db     639 FERELW---LSGIANFFYRDSMPTRHGFRHISGGYALGITATTPAEDQLTFA--FCQLFA 693

Qy     747 RYAKNNVSSKN---------YSCQGEMLFSLQEGFL-------------LTKLVGL---- 780
Db     694 R-DRNHITGKNHGDTYGASLYFHHTEGLFDI-ANFLWGKATRAPWVLSEISQIIPLSFDA 751

Qy     781 ---YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHIL--TAPFLGA 835
Db     752 KFSYLHTDNHMKTYYT--DNSIIKGSWRNDAFCADLGASLPF-VISVPYLLKEVEPFVKV 808

Qy     836 LGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVL- 894
Db     809 QYIYAHQQDFYERHAEGRAFN-KSELINVEIPIGVT---FERDSKSEKGTYDLTLMYILD 864

Qy     895 -YRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTF 953
Db     865 AYRRNPKCQTSLIASDANWMAYGTNLARQGFSVRAANHFQVNPHMEIFGQF-AFEVRSSS 923

Qy     954 CNYLNGEIALRF 965
Db     924 RNY-NTNLGSKF 934
```

RESULT 15
C86546
polymorphic outer membrane protein G family [imported] - Chlamydophila pneumoniae (st
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 15-Jun-2001
C;Accession: C86546
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;
Nucleic Acids Res. 28, 2311-2314, 2000
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;Reference number: A86491; MUID:20330349; PMID:10871362
A;Accession: C86546
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-936 <STO>
A;Cross-references: GB:BA000008; NID:g8978817; PIDN:BAA98653.1; GSPDB:GN00142
A;Experimental source: strain J138
C;Genetics:
A;Gene: pmp_7
C;Superfamily: Chlamydophila pneumoniae polymorphic outer membrane protein G

Query Match 12.7%; Score 644; DB 2; Length 936;
Best Local Similarity 27.0%; Pred. No. 1.5e-31;
Matches 262; Conservative 146; Mismatches 392; Indels 172; Gaps 44;

```
Qy      75 GAAVTITDYLSFFDTQKEGIYFAKNLTPESGGAIGY------------ASPNSPTVEIR 121
Db      54 GTTYSLLSDVSFQNAGALGIPLASGCFLEAGGDLTFQGNQHALKFAFINAGSSAGTVAST 113

Qy     122 DTIGPVIFENN-------TCCRPFTSSNPNAAVNKIREGGAIHAQNLYINHNHDVVGFMK 174
Db     114 SAADKNLLFNDFSRLSIISCPSLLLSPTGQCALKSV--------GNLSLTGNSQII-FTQ 164

Qy     175 NFSYVRGGAISTANTFVVSENQSCFLFMDNICIQTNTAGKGGAIYA-GTSNSFESNNCDL 233
Db     165 NFSSDNGGVINTKN-FLLSGTSQFASFSRN---QAFTGKQGGVVYATGTITIENSPGIVS 220

Qy     234 FFINNACCAGGAIFS-PICSLTGNRGNIVFYNNRCFKNVETASSEASDGGAIKVTT---R 289
Db     221 FSQNLAKGSGGALYSTDNCSITDN-FQVIFDGNSAWEAAQ------AQGGAICCTTTDKT 273

Qy     290 LDVTGNRGRIFFSDNITKNYGGAIYAPVVTLVDNGPTYFINNIANNK-----GGAIYIDG 344
Db     274 VTLTGNK-NLSFTNNTALTYGGAISGLKVSISAGGPTLFQSNISGSSAGQGGGGAINIAS 332

Qy     345 TSNSKISADRHAIIFNENIVTNVTNANGTSTSANPPRRNAITVASSSGEILLGAGSSQNL 404
Db     333 AGELALSATSGDITFNNNQVTN----GSTST------RNAINIIDTAKVTSIRAATGQSI 382

Qy     405 IFYDPIEVSNAGVSVS----------FNKEADQTGSVVFSGATVNSAD-FHQRNLQTKTP 453
Db     383 YFYDPI--TNPGTAASTDTLNLNLADANSEIEYGGAIVFSGEKLSPTEKAIAANVTSTIR 440

Qy     454 APLTLSNGFLCIEDHAQLTVNRFTQTGGVVSLGNGAVLSCYKNGAGNSASNASITLKHIG 513
Db     441 QPAVLARGDLVLRDGVTVTFKDLTQSPGSRILMDG--------GTTLSAKEANLSLNGLA 492

Qy     514 LNLSSILKSGAEIPLLWVEPTNNSNNYTADTAATFSLSDVKLSLIDDYGNSPYESTDLTH 573
Db     493 VNLSSL------------DGTNKAALKTEAADKNISLSGT-IALIDTEG-SFYENHNLKS 538

Qy     574 ALSSQPMLSISEASDN-QLRSDDMDFSGLNVP--HYGWQGLWTWGWAKTQDPEPASSATI 630
Db     539 A-STYPLLELTTAGANGTITLGALSTLTLQEPETHYGYQGNWQLSWAN------ATSSKI 591

Qy     631 TDPQKANRFHRTLLLTWLPAGYVPSPKHRSPLIANTLWGNMLLATESLKNSAELTPSDHP 690
Db     592 GS------------INWTRTGYIPSPERKSNLPLNSLWGN-FIDIRSINQLIETKSSGEP 638
```

```
Qy     691 F----WGITGGGLGMMVYQDPRENHPGFHMRSSGYSAGMIAGQTHTFSLKFSQTYTKLNE 746
           |    |     |:    |:|      ||   | ||: |: |       | |:  : :|
Db     639 FERELW---LSGIANFFYRDSMPTRHGFRHISGGYALGITATTPAEDQLTFA--FCQLFA 693

Qy     747 RYAKNNVSSKN---------YSCQGEMLFSLQEGFL-------------LTKLVGL---- 780
           |  :|::: ||         |    | || :   ||             :::::|
Db     694 R-DRNHITGKNHGDTYGASLYFHHTEGLFDI-ANFLWGKATRAPWVLSEISQIIPLSFDA 751

Qy     781 ---YSYGDHNCHHFYTQGENLTSQGTFRSQTMGGAVFFDLPMKPFGSTHIL--TAPFLGA 835
              | : |::   :||  :|   :|::|:      :   ||       ::|    ||:
Db     752 KFSYLHTDNHMKTYYT--DNSIIKGSWRNDAFCADLGASLPF-VISVPYLLKEVEPFVKV 808

Qy     836 LGIYSSLSHFTEVGAYPRSFSTKTPLINVLVPIGVKGSFMNATQRPQAWTVELAYQPVL- 894
             ||:    | |  | :|:|: |: ||||  :||||        : :  | :|    :|
Db     809 QYIYAHQQDFYERYAEGRAFN-KSELINVEIPIGVT---FERDSKSEKGTYDLTLMYILD 864

Qy     895 -YRQEPGIATQLLASKGIWFGSGSPSSRHAMSYKISQQTQPLSWLTLHFQYHGFYSSSTF 953
            ||: |   | |:||   |   |:  :|   | : :   |    : :  |:  |   |:
Db     865 AYRRNPKCQTSLIASDANWMAYGTNLARQGFSVRAANHFQVNPHMEIFGQF-AFEVRSSS 923

Qy     954 CNYLNGEIALRF 965
             || |  :  :|
Db     924 RNY-NTNLGSKF 934
```

Search completed: February 4, 2003, 20:58:52
Job time : 56 secs